Cuk-H02[illegible]-29-P010932

THE ACCOUNTING REVIEW

VOLUME 77, SUPPLEMENT 2002
QUALITY OF EARNINGS CONFERENCE

The Balance Sheet as an Earnings Management Constraint
Jan Barton and Paul J. Simko
DISCUSSION: Mark L. DeFond

The Quality of Accruals and Earnings: The Role of Accrual Estimation Errors
Patricia M. Dechow and Ilia D. Dichev
DISCUSSION: Maureen F. McNichols

The Relation between Auditors' Fees for Nonaudit Services and Earnings Management
Richard M. Frankel, Marilyn F. Johnson, and Karen K. Nelson
DISCUSSION: William R. Kinney, Jr. and Robert Libby

Empirical Evidence on the Evolution of International Earnings
Judy Land and Mark H. Lang
DISCUSSION: David A. Guenther

Evaluating Financial Reporting Quality: The Effects of Financial Expertise vs. Financial Literacy
Linda McDaniel, Roger D. Martin, and Laureen A. Maines
DISCUSSION: Mark E. Peecher

Evidence from Auditors about Managers' and Auditors' Earnings Management Decisions
Mark W. Nelson, John A. Elliott, and Robin L. Tarpley
DISCUSSION: Michael Gibbins

A QUARTERLY JOURNAL OF THE
AMERICAN ACCOUNTING ASSOCIATION

AMERICAN ACCOUNTING ASSOCIATION

The members of the 2002–2003 Executive Committee of the Association are as follows:

President	G. Peter Wilson, Boston College
President-Elect	William L. Felix, Jr., The University of Arizona
Vice President	William E. (Bill) McCarthy, Michigan State University
Vice President–Education	Kevin D. Stocks, Brigham Young University
Vice President–Finance	Judy D. Rayburn, University of Minnesota
Vice President–International	Judy S. L. Tsui, The Hong Kong Polytechnic University
Vice President–Professional Relations	Samuel A. Vitkoski, BDO Seidman
Vice President–Publications	Lawrence A. Tomassini, The Ohio State University
Vice President–Research	Zoe-Vonna Palmrose, University of Southern California
Vice President–Sections and Regions	Joanna L. Ho, University of California, Irvine
Past President	Joel S. Demski, University of Florida

The American Accounting Association was founded in 1916 as the American Association of University Instructors in Accounting. Its present name was adopted in 1936. The Association is a voluntary organization of persons interested in accounting education, research, and practice. Full membership entitles one to receive *Accounting Education News*, a selection of one, two, or all three of the Association-wide journals (*The Accounting Review, Accounting Horizons,* and *Issues in Accounting Education*) and the opportunity to purchase Association publications at a discount.

Membership in the Association is available at the following annual rates:

Full membership:
The membership dues are based on the selection of one, two, or all three of the following Association-wide journals: *The Accounting Review, Accounting Horizons,* and *Issues in Accounting Education.*
(This does not include Section journals.)

Select one Association-wide journal	U.S. $85.00
Select two Association-wide journals	U.S. $95.00
All three Association-wide journals	U.S. $105.00

Associate membership for students residing in the United States and Canada:

Select one Association-wide journal	U.S. $20.00
Select two Association-wide journals	U.S. $30.00
All three Association-wide journals	U.S. $40.00

Electronic journal option available for an additional U.S. $20.00 to above-mentioned full-membership and associate-membership dues.

In addition, full members may belong to one or more Sections: Accounting, Behavior and Organizations; Accounting Programs Leadership Group; American Taxation Association; Artificial Intelligence/Emerging Technologies; Auditing; Financial Accounting and Reporting; Gender Issues in Accounting; Government and Nonprofit; Information Systems; International Accounting; Management Accounting; Public Interest; Teaching and Curriculum; and Two-Year College. Each Section assesses its own annual dues, which are collected by the Association.

Editorial correspondence and all manuscripts should be sent to Professor Linda Smith Bamber, Editor, *The Accounting Review*, Tull School of Accounting, Terry College of Business, University of Georgia, 255 Brooks Hall, Athens, GA 30602-6252.

The submission fee of $75.00 for members of the AAA or $100.00 for nonmembers must accompany manuscripts when submitted. Books for review should be sent to Jane L. Reimers, College of Business, Florida State University, Tallahassee, FL 32306-1042. Request for reprint permission, subscriptions, advertisements, and dues should be sent to American Accounting Association, 5717 Bessie Drive, Sarasota, FL 34233-2399 • Telephone: (941) 921-7747 • Fax: (941) 923-4093 • Email: Office@aaahq.org. Notice of change of address should be mailed at least eight weeks in advance to ensure undelayed receipt of *The Accounting Review*.

The Accounting Review (ISSN 0001-4826) is published four times a year in January, April, July, and October by the American Accounting Association, 5717 Bessie Drive, Sarasota, Florida 34233-2399. General subscriptions are available at the annual rate of U.S. $170; membership subscription rate is $25 with discounts available; and emeritus subscription rate is $8.75. Periodical postage paid at Sarasota, Florida and additional mailing offices. **POSTMASTER:** Send address change to ***The Accounting Review,*** 5717 Bessie Drive, Sarasota, Florida 34233-2399.

© 2002 American Accounting Association. All rights reserved. Printed in the United States of America.

P-10,932

657.05
AC57R

THE ACCOUNTING REVIEW

A QUARTERLY JOURNAL OF THE
AMERICAN ACCOUNTING ASSOCIATION

QUALITY OF EARNINGS CONFERENCE

Editor
Katherine Schipper

29

VOLUME 77 SUPPLEMENT 2002

The Balance Sheet as an Earnings Management Constraint 1
Jan Barton and Paul J. Simko
DISCUSSION: Mark L. DeFond

The Quality of Accruals and Earnings: The Role of Accrual Estimation Errors 35
Patricia M. Dechow and Ilia D. Dichev
DISCUSSION: Maureen F. McNichols

The Relation between Auditors' Fees for Nonaudit Services and Earnings Management 71
Richard M. Frankel, Marilyn F. Johnson, and Karen K. Nelson
DISCUSSION: William R. Kinney, Jr. and Robert Libby

Empirical Evidence on the Evolution of International Earnings 115
Judy Land and Mark H. Lang
DISCUSSION: David A. Guenther

Evaluating Financial Reporting Quality: The Effects of Financial Expertise vs. Financial Literacy 139
Linda McDaniel, Roger D. Martin, and Laureen A. Maines
DISCUSSION: Mark E. Peecher

Evidence from Auditors about Managers' and Auditors' Earnings Management Decisions 175
Mark W. Nelson, John A. Elliott, and Robin L. Tarpley
DISCUSSION: Michael Gibbins

Editorial Staff of *The Accounting Review*
2001–2002

SUPPLEMENT EDITOR
Katherine Schipper, Federal Accounting Standards Board
EDITOR
Linda Smith Bamber, University of Georgia
CONSULTING EDITOR
E. Michael Bamber, University of Georgia
ASSOCIATE EDITORS
Greg Clinch, Melbourne Business School
Raffi J. Indjejikian, University of Michigan
Steven J. Kachelmeier, The University of Texas at Austin
Kenneth J. Klassen, University of Waterloo
Marlys Gascho Lipe, University of Oklahoma
Joan L. Luft, Michigan State University
Mark W. Nelson, Cornell University
Grace Pownall, Emory University
Richard C. Sansing, Dartmouth College
Daniel B. Thornton, Queen's University
Senyo Tse, The University of Texas at Austin
BOOK REVIEW EDITOR
Jane L. Reimers, Florida State University

American Accounting Association

INTERIM EXECUTIVE DIRECTOR
Tracey E. Sutherland
COMMUNICATIONS DEPARTMENT
Beverly Harrelson
Diane Hazard
Kathy Casper
Helen Griffith
Carol Stanton
Judy Cothern
Jan Kovarik (subcontracted)
OTHER AAA STAFF
Debora Gardner
Roland LaTulip
Patricia A. Hoban
Deirdre Harris
Mary Cole
James Szelminski
Linda Watterworth
Mary Cole
Barbara Brady
Angela Kasapakis

Editorial Advisory and Review Board

Ashiq Ali, The University of Arizona
Shannon W. Anderson, Rice University
Benjamin C. Ayers, University of Georgia
Frances L. Ayres, University of Oklahoma
William R. Baber, The George Washington University
Stephen P. Baginski, University of Georgia
Ramamurthy Balakrishnan, The University of Iowa
Orie E. Barron, Pennsylvania State University
Jan Barton, Emory University
Anne L. Beatty, Pennsylvania State University
Jacob G. Birnberg, University of Pittsburgh
Robert J. Bloomfield, Cornell University
Robert M. Bowen, University of Washington
Lawrence D. Brown, Georgia State University
David C. Burgstahler, University of Washington
Brian J. Bushee, University of Pennsylvania
Robert M. Bushman, The University of North Carolina at Chapel Hill
Carolyn M. Callahan, University of Arkansas
Youngsoon Susan Cheon, Chung-Ang University
Kevin C. W. Chen, Hong Kong University of Science & Technology
Chee W. Chow, San Diego State University
C. Bryan Cloyd, University of Illinois at Urbana–Champaign
Paul A. Copley, University of Georgia
Patricia M. Dechow, University of Michigan
Mark L. DeFond, University of Southern California
Ilia D. Dichev, University of Michigan
Sunil Dutta, University of California, Berkeley
Pieter T. Elgers, University of Massachusetts
Merle M. Erickson, University of Chicago
John (Harry) Evans III, University of Pittsburgh
Gerald A. Feltham, University of British Columbia
Paul E. Fischer, Pennsylvania State University
Joseph G. Fisher, Indiana University
Jere R. Francis, University of Missouri–Columbia
O. Finley Graves, Kansas State University
Wayne R. Guay, University of Pennsylvania
David A. Guenther, University of Colorado at Boulder
Mahendra Gupta, Washington University
Susan F. Haka, Michigan State University
John R. M. Hand, The University of North Carolina at Chapel Hill
Donald R. Herrman, Oregon State University
D. Eric Hirst, The University of Texas at Austin
Chris E. Hogan, Southern Methodist University
Patrick E. Hopkins, Indiana University
K. E. (Skip) Hughes II, Louisiana State University
Christopher D. Ittner, University of Pennsylvania
Debra C. Jeter, Vanderbilt University
James J. Jiambalvo, University of Washington
Deen Kemsley, Columbia University
S. Jane Kennedy, University of Washington
William R. Kinney, Jr., The University of Texas at Austin
Lisa L. Koonce, The University of Texas at Austin
William N. Lanen, University of Michigan
Charles M. C. Lee, Cornell University
Robert Libby, Cornell University
Robert C. Lipe, University of Oklahoma
Russell J. Lundholm, University of Michigan

Joseph Magliolo III, Southern Methodist University
Laureen A. Maines, Indiana University
Ella Mae Matsumura, University of Wisconsin–Madison
Steven R. Matsunaga, University of Oregon
Linda S. McDaniel, University of Kentucky
Maureen F. McNichols, Stanford University
Richard R. Mendenhall, University of Notre Dame
Lillian F. Mills, The University of Arizona
James N. Myers, University of Illinois at Urbana–Champaign
Patricia C. O'Brien, University of Waterloo
Mark E. Peecher, University of Illinois at Urbana–Champaign
Kathy R. Petroni, Michigan State University
Jamie H. Pratt, Indiana University
Robert J. Ramsay, University of Kentucky
Madhav V. Rajan, Stanford University
Stephen G. Ryan, New York University
Katherine Schipper, Financial Accounting Standards Board
D. Shores, University of Washington
Paul J. Simko, University of Virginia
Richard G. Sloan, University of Michigan
J. Reed Smith, SUNY at Buffalo
Geoffrey B. Sprinkle, Indiana University
Thomas L. Stober, University of Notre Dame
Steve G. Sutton, University of Connecticut/University of Melbourne
Hun-Tong Tan, Nanyang Technological University
Siew Hong Teoh, The Ohio State University
Lynda Thoman, Purdue University
Richard Tubbs, The University of Iowa
Igor Vaysman, INSEAD
Robert E. Verrecchia, University of Pennsylvania
Martha L. Wartick, University of Northern Iowa
Gregory B. Waymire, Emory Univeristy
Hong Xie, University of Illinois at Urbana–Champaign
Teri L. Yohn, Georgetown University

PAST EDITORS

1926–1929	William A. Paton
1929–1943	Eric L. Kohler
1944–1947	A. C. Littleton
1948–1949	Robert L. Dixon
1950–1959	Frank P. Smith
1960–1962	Robert K. Mautz
1963–1964	Lawrence L. Vance
1965–1967	Wendell Trumbull
1968–1970	Charles H. Griffin
1971–1972	Eldon S. Hendriksen
1973–1975	Thomas F. Keller
1976–1978	Don T. DeCoster
1978–1983	Stephen A. Zeff
1983–1987	Gary L. Sundem
1987–1990	William R. Kinney, Jr.
1990–1994	A. Rashad Abdel–khalik
1994–1997	Robert P. Magee
1997–2000	Gerald L. Salamon

THE ACCOUNTING REVIEW
Vol. 77
Supplement 2002
pp. 1–27

The Balance Sheet as an Earnings Management Constraint

Jan Barton
Emory University
Paul J. Simko
University of Virginia

ABSTRACT: The balance sheet accumulates the effects of previous accounting choices, so the level of net assets partly reflects the extent of previous earnings management. We predict that managers' ability to optimistically bias earnings decreases with the extent to which the balance sheet overstates net assets relative to a neutral application of GAAP. To test this prediction, we examine the likelihood of reporting various earnings surprises for 3,649 firms during 1993–1999. Consistent with our prediction, we find that the likelihood of reporting larger positive or smaller negative earnings surprises decreases with our proxy for overstated net asset values.

Keywords: *articulation between income statement and balance sheet; earnings management constraints; measurement and recognition guidance; earnings surprises; analysts' forecasts.*

Data Availability: *Data are available from sources identified in the paper.*

I. INTRODUCTION

We examine the extent to which Generally Accepted Accounting Principles (GAAP) and their implementation guidelines constrain managers from optimistically biasing earnings. Because the balance sheet accumulates the effects of previous accounting choices, the level of net assets partly reflects the extent of previous earnings management. We predict that managers' ability to optimistically bias earnings decreases with the extent to which net assets are already overstated on the balance sheet.[1] To test our prediction, we focus on a particularly strong earnings management

---

[1] We use the term "overstated" to describe net assets measured and recorded at values exceeding those that would have been reported under a neutral application of GAAP, without implying a violation of GAAP.

We appreciate the helpful comments of Katherine Schipper (the editor), an anonymous referee, Linda Bamber, George Benston, Walt Blacconiere, Paul Irvine, Connie Kertz, Steven Schmitt, Greg Waymire, and seminar participants at The University of Alabama, University of Arkansas, Emory University, Georgia State University, Indiana University, University of Virginia, *The Accounting Review* Conference on Quality of Earnings, and The University of Texas at Austin's 2002 BMAS Conference. We also thank Ron Harris and Xiao Lan Wang for research assistance, the Goizueta Business School for generous funding, and I/B/E/S International Inc. for access to data on analysts' earnings forecasts.

incentive—meeting or beating analysts' earnings expectations. Meeting or beating expectations is not always the result of optimistic bias in financial reporting, so our empirical analyses control for other reasons behind an earnings surprise. Our analyses suggest that managers are less likely to report a predetermined earnings surprise by optimistically biasing earnings when their firms' net assets are overstated.

The articulation between the income statement and the balance sheet ensures that accruals reflected in earnings also are reflected in net assets. Therefore, an optimistic bias in earnings implies net assets measured and recorded temporarily at values exceeding those based on a neutral application of GAAP. Managers' generous assumptions about recognition and measurement in one period reduce their ability to make equally generous assumptions in later periods, if managers want to stay within the guidance provided by accounting regulators and professional groups.[2] Therefore, managers' ability to optimistically bias earnings decreases with the extent to which net assets are already overstated.

We use the beginning balance of net operating assets relative to sales as a proxy for managers' previous biased reporting choices. This proxy is consistent with overstated net assets being less efficient at generating a given level of sales, all else equal. If this proxy is valid, then firms with larger levels of net operating assets relative to sales will have reported larger cumulative levels of income-increasing accruals in the past. In our sample of 3,649 nonfinancial, nonregulated firms during 1993–1999, we find that firms with larger levels of net operating assets (relative to sales) reported larger cumulative levels of abnormal accruals in the previous 20 quarters, consistent with prior income-increasing earnings management leading to overstated net assets.

We then examine the association between our sample firms' quarterly earnings surprises and the beginning balance of net operating assets relative to sales. We focus on earnings surprises because "[p]erhaps the single most important cause [of earnings management] is the pressure imposed on management to meet analysts' earnings projections" (Johnson 1999). Investors have become especially unforgiving of firms that fail to meet earnings expectations (Skinner and Sloan 2001).[3] Previous research documents an increasing incidence of reported earnings that just meet or slightly beat analysts' earnings forecasts (Brown 2001; Burgstahler and Eames 2001; Matsumoto 2002), an outcome that is apparently achieved in part through earnings management (Payne and Robb 2000; Burgstahler and Eames 2001; Matsumoto 2002). If missing expectations is costly and managing earnings toward expectations is common, then why do some managers miss expectations even by a small amount? In fact, why do they meet expectations rather than beat them altogether? We suggest some managers do so because, all else equal, they have limited discretion to repeatedly bias earnings upward. Specifically, we predict that managers' ability to report larger positive or smaller negative earnings surprises decreases with the extent to which net assets are already overstated on the balance sheet.

We model the level of earnings surprise as a function of managers' previous recognition and measurement decisions, as reflected in the beginning balance of net operating assets relative to sales. Our model controls for other constraints on earnings management, managerial incentives to meet or beat analysts' forecasts, and firm performance and size. We estimate the model using a generalized ordered logit regression on 35,950 quarterly earnings surprises for our sample. Consistent with our prediction, we find that firms with larger beginning balances of net operating assets relative to sales are less likely to report a predetermined earnings surprise. These findings are robust to using alternative econometric techniques, separating net operating assets into their current and long-term components, controlling for industry and fiscal-quarter effects, using alternative definitions of earning surprises, and excluding loss firms or firms reporting special items.

---

[2] Departing from GAAP may result in substantial costs to managers and their firms (Dechow et al. 1996).

[3] For example, Compaq Computer, Hewlett-Packard, Merck, Sears, Starbucks, and UAL all missed their consensus analyst forecasts by 1¢ per share at least once during 1993–1999, and suffered an average size-adjusted cumulative abnormal return of about −12 percent over the three days surrounding their earnings announcements.

Our paper extends research on earnings management in three ways. First, while we do not measure directly the amount of financial reporting discretion actually available to managers, our approach provides a basis for doing so in future research. Second, prior research examines how managers trade off financial reporting discretion available across several balance sheet accounts when meeting various objectives (e.g., Beatty et al. 1995; Hunt et al. 1996; Gaver and Paterson 1999). We extend this literature by providing evidence that the balance sheet accumulates the effects of prior optimism in financial reporting, reducing managers' ability to optimistically bias future earnings. Therefore, consistent with Hunt et al. (1996), our findings suggest that managers likely trade off available financial reporting discretion over time. Last, prior research provides evidence that managing earnings to meet expectations is widespread (Degeorge et al. 1999; Payne and Robb 2000; Abarbanell and Lehavy 2001a; Burgstahler and Eames 2001; Matsumoto 2002), but assumes that the ability to manage earnings is random or even constant across firms. Our results suggest that the ability to report a predetermined earnings surprise varies across firms as a function of previous optimism in financial reporting.

Section II develops our main hypothesis. Section III describes our research design. Specifically, this section explains the sample selection criteria, reports the frequencies and market reactions associated with various levels of earnings surprise, describes the characteristics of sample firms, discusses our proxy for overstated net asset values, and develops our regression model. Section IV describes our main empirical results and sensitivity tests. Finally, Section V offers some concluding remarks.

II. HYPOTHESIS

The accruals basis of accounting requires managers to report the effects of economic transactions based on expected cash realizations (SFAC No. 1, FASB 1978; SFAC No. 6, FASB 1985). Although GAAP requires managers to make numerous judgments and assumptions to report their firms' performance under accrual accounting, managers do not have unlimited discretion in doing so. For example, accounting regulators like the Securities and Exchange Commission and the Financial Accounting Standards Board, along with professional groups like the American Institute of Certified Public Accountants, provide detailed guidelines for implementing measurement, recognition, and disclosure rules. Auditors and the courts help enforce these guidelines.

Because of the reversing nature of accrual accounting, managers' biased estimates and judgments in one period reduce their ability to make similarly biased estimates and judgments in subsequent periods (Hunt et al. 1996; Abarbanell and Lehavy 2001a). The articulation between the income statement and the balance sheet ensures that biased assumptions reflected in earnings are also reflected in net asset values. Therefore, we predict that managers' ability to optimistically bias earnings decreases with the extent to which net assets are already overstated on the balance sheet, relative to what their values would have been under a neutral implementation of GAAP.

Prior research shows that overstated assets and revenues are the most common reason for restatements of financial statements (Palmrose et al. 2001), Securities and Exchange Commission's accounting enforcement actions (Feroz et al. 1991; Dechow et al. 1996), and lawsuits against auditors (St. Pierre and Anderson 1984; Lys and Watts 1994). Prior research also shows that managers may understate assets temporarily to regain the ability to optimistically bias earnings in future periods (Moehrle 2002).

Managers have various incentives to bias earnings (see Fields et al. 2001). However, we focus on one particularly strong incentive—meeting or beating analysts' earnings forecasts. Dechow and Skinner (2000) argue that the bull market of the late 1990s and the increased use of stock options for managerial compensation have increased managers' incentives to manipulate earnings to preserve high stock valuations. Consistent with this argument, Barth et al. (1999) provide evidence that firms with patterns of increasing earnings have higher price-to-earnings ratios, and that these ratios decline significantly when firms break the pattern by reporting earnings decreases. Moreover, Skinner and Sloan (2001) report significant negative returns for firms that fail to meet analysts' earnings forecasts. Finally, evidence also suggests widespread earnings management to meet or beat analysts' forecasts (e.g., Payne and Robb 2000; Burgstahler and Eames 2001; Matsumoto 2002).

Based on these findings, we predict that managers will use available financial reporting discretion to report higher levels of earnings surprises, all else equal. Therefore, we test the following hypothesis:

> **H_a:** The likelihood of reporting larger positive or smaller negative earnings surprises decreases with the extent to which net assets are already overstated on the balance sheet.

III. RESEARCH DESIGN

Sample

Our initial sample contains all firm-quarters on both the Compustat and I/B/E/S databases with complete data for our test variables during 1993–1999. We consider this period because pressure to meet analysts' forecasts appears to be a recent phenomenon (Levitt 1998; Brown 2001; Matsumoto 2002), and because Abarbanell and Lehavy (2001b) suggest that differences between pre- and post-1993 I/B/E/S forecast errors may be due to changes in the way I/B/E/S calculates actual earnings per share (EPS). We exclude utilities and financial services firms (two-digit SIC codes 49 and 60–67) because they are subject to regulatory requirements that would unnecessarily complicate our research design. Our final sample consists of 35,950 quarterly observations for 3,649 firms.

We define an earnings surprise (SURPRISE) as the I/B/E/S actual EPS for quarter t less the consensus analyst EPS forecast for quarter t, both rounded to the nearest penny. The consensus forecast is the average of analysts' most recent forecasts of EPS for quarter t available on I/B/E/S prior to the earnings announcement for quarter t (Brown and Kim 1991). To keep our analyses tractable, we combine firms missing the consensus forecast by 5¢ or more (i.e., SURPRISE ≤ -5¢) into one category and firms beating the forecast by 5¢ or more (i.e., SURPRISE ≥ 5¢) into another. SURPRISE thus consists of 11 categories coded sequentially from −5 to 5.

Table 1 reports characteristics of quarterly earnings surprises for our sample. Here and in our remaining analyses, we winsorize all variables at the upper and lower one percentile of their distributions. Across levels of SURPRISE, Panel A of Table 1 reports the means of actual EPS, consensus EPS forecast, and forecast error (i.e., actual EPS less consensus EPS forecast) as a percentage of the consensus EPS forecast. Panel A shows these three variables increasing in the level of SURPRISE. The panel also shows that firms missing (beating) earnings expectations by 1¢ per share reported EPS averaging about 6.3 percent lower (6.7 percent higher) than the consensus forecast. On the other hand, firms missing (beating) earnings expectations by 5¢ or more reported EPS averaging about 31 percent lower (20.4 percent higher) than the consensus forecast.

TABLE 1
Characteristics of Quarterly Earnings Surprises

Panel A: Mean of Actual Earnings Per Share (EPS), Consensus Analysts' Forecasts, and Earnings Surprise as a Percentage of Forecasts

SURPRISE	*n*	*Actual EPS*	*Consensus EPS Forecast*	*SURPRISE / Consensus EPS Forecast*
≤ −5¢	5,054	2¢	17¢	−31.0%
−4	871	16	20	−16.1
−3	1,196	17	20	−13.5
−2	1,707	20	22	−8.3
−1	2,811	23	24	−6.3
0	8,426	22	22	0.6
1	5,163	25	24	6.7
2	3,050	27	25	10.7
3	1,927	28	25	12.2
4	1,236	28	24	13.9
≥ 5	4,509	38	28	20.4
All firm-quarters	35,950	22¢	23¢	−0.4%

Panel B: Annual Frequency of Firm-Quarters Reporting a Given Earnings Surprise

	Fiscal Year							
SURPRISE	*1993*	*1994*	*1995*	*1996*	*1997*	*1998*	*1999*	*Total*
≤ −5¢	16.4%	13.8%	17.1%	14.0%	13.3%	13.3%	11.9%	14.1%
−4	3.7	2.7	2.8	2.3	1.8	2.4	1.9	2.4
−3	4.2	3.9	3.7	3.1	2.8	3.6	2.5	3.3
−2	6.2	5.3	4.8	5.4	4.6	4.2	3.4	4.8
−1	9.8	8.6	7.7	7.8	7.6	7.7	6.4	7.8
0	25.4	24.5	24.0	24.8	24.1	22.5	20.1	23.4
1	10.4	13.4	13.7	13.7	15.6	15.9	16.0	14.4
2	6.6	7.6	7.2	8.5	9.1	9.3	10.0	8.5
3	4.2	4.7	4.8	4.9	6.2	5.6	6.4	5.4
4	2.4	3.3	3.1	3.3	3.2	3.5	4.8	3.4
≥ 5	10.7	12.2	11.1	12.2	11.7	12.0	16.6	12.5
Total	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
n	3,806	4,371	4,743	5,305	5,757	6,175	5,793	35,950

(*Continued on next page*)

TABLE 1 (Continued)

Panel C: Mean Three-Day Cumulative Abnormal Return (CAR) around Quarterly Earnings Announcement Date[a]

SURPRISE	*n*	*CAR*	*t-Statistic for H_0: CAR = 0*	*CAR / SURPRISE*
≤ −5¢	4,877	−2.81%	−25.02 ***	−0.27%
−4	846	−1.84	−7.60 ***	−0.46
−3	1,162	−1.98	−9.24 ***	−0.66
−2	1,670	−1.41	−8.44 ***	−0.71
−1	2,740	−1.16	−8.95 ***	−1.16
0	8,221	−0.01	−0.12	NA
1	4,983	1.04	10.49 ***	1.04
2	2,930	2.00	15.90 ***	1.00
3	1,857	2.17	13.47 ***	0.72
4	1,177	2.81	13.28 ***	0.70
≥ 5	4,278	3.18	28.81 ***	0.42
All firm-quarters	34,741	0.26%	6.51 ***	0.19%

*** denotes significant at the 0.01 level, based on two-tailed tests.

[a] We calculate the cumulative abnormal return (CAR) as the with-dividend return over the three days surrounding the earnings announcement for quarter t, less the return over the same period on a size-matched portfolio. To calculate the size-matched portfolio's return, we allocate all firm-quarters to decile portfolios based on market value of common equity at the beginning of quarter t, and then we calculate a value-weighted return for each portfolio for the three-day period surrounding each earnings announcement.

The sample consists of 35,950 firm-quarters pertaining to 3,649 firms included in both the Compustat and I/B/E/S databases with complete data for our primary tests over 1993−1999, excluding utilities and financial services firms (two-digit SIC codes 49 and 60−67). Panel C also excludes 1,209 firm-quarters with incomplete data to calculate abnormal returns. SURPRISE is the I/B/E/S actual EPS for quarter t less the consensus forecast for quarter t, both rounded to the nearest penny. The consensus forecast is the mean of analysts' most recent EPS forecasts for quarter t available on I/B/E/S prior to the earnings announcement for quarter t. In Panels A and C, n denotes the number of firm-quarters in each category of SURPRISE; in Panel B, n denotes the number of firm-quarters per fiscal year.

Panel B of Table 1 shows that the frequency of firms beating earnings expectations increased over the sample period, consistent with findings reported in Brown (2001) and Matsumoto (2002). A Pearson χ^2 test (not tabulated) rejects the null of no association between SURPRISE and fiscal year ($\chi^2_{(60\ df)} = 583.73$, $p < 0.01$).

Panel C of Table 1, which presents mean three-day cumulative abnormal returns (CAR) around the earnings announcement date across levels of SURPRISE, confirms the well-known positive association between CARs and earnings surprises.[4] The mean CAR for firms meeting expectations is −0.01 percent, statistically insignificant at conventional levels. The last column of the panel shows that the average proportionate loss in equity value is more severe for firms missing expectations by 1¢ than for firms missing by a larger amount ($t = -5.39$, $p < 0.01$, not tabulated). For example, an EPS surprise of −1¢ is associated with a mean CAR of −1.16 percent, whereas an EPS surprise of −4¢ (four times larger) is associated with a mean CAR of −1.84 percent, about −0.46 percent per penny of surprise.

---

[4] As in Skinner and Sloan (2001), our CARs are based on size-matched portfolios. Specifically, we calculate each firm-quarter's CAR as the with-dividend return over the three days surrounding the earnings announcement, less the return on a size-matched portfolio over the same period. To calculate the size-matched portfolio's return, we allocate all firm-quarters to decile portfolios based on market values of common equity at the beginning of the quarter. We then calculate a value-weighted return on each portfolio around each earnings announcement date. We are unable to obtain returns for 1,209 firm-quarters (3.4 percent of our sample).

Table 2 reports descriptive statistics for selected firm characteristics. On average, our sample firms have a market capitalization of $2.3 billion, total assets of $1.8 billion, and quarterly sales of $471 million. About 44.1 percent of the firms are traded on NASDAQ, 32.5 percent of their assets are fixed (i.e., consisting of property, plant, and equipment, net of accumulated depreciation), and 61.7 percent of their total assets are financed through debt. Although the average net income is $24.5 million per quarter, 18.6 percent of the firm-quarters report losses. The average quarterly operating cash flow is $102 million, about four times larger than average net income.

The last column of Table 2 reports Spearman rank correlations between the selected firm characteristics and SURPRISE. Although significant at conventional levels, most correlations are less than 0.20 in absolute terms. The correlations suggest that firms reporting higher (i.e., larger positive or smaller negative) levels of SURPRISE are larger in terms of market capitalization, total assets, sales, net income, and operating cash flows. These firms also are less levered, and less likely both to report losses and to be traded on NASDAQ.

Measuring Overstatement in Net Asset Values

We use the beginning balance of net operating assets relative to sales (NOA) as a proxy for bias in the implementation of GAAP measurement and recognition guidelines. We define net operating assets as shareholders' equity less cash and marketable securities, plus total debt. Therefore, NOA consists mainly of accrual-based measures of net assets used in operations. This proxy is consistent with the assumption that overstated net assets are less efficient at generating a given level of sales, all else equal. However, reported sales may also reflect managers' biased implementation of GAAP. Unlike net asset amounts, sales reflect reporting bias exerted only during the current period. If sales also are overstated, then using NOA will work against finding support for H_a.

To provide evidence that NOA captures previous optimism in financial reporting, we test whether firms with larger NOA also reported larger cumulative levels of income-increasing accruals in the past. To this end, we first estimate abnormal accruals for firm i in quarter t using the residual from the following regression, estimated by two-digit SIC code and fiscal year using 1988–1999 data (see Jones 1991; Dechow et al. 1995; Han and Wang 1998):

$$\begin{aligned} TOT\_ACC_{it}/TA_{it-1} = {} & \phi_1(1/TA_{it-1}) + \phi_2[(\Delta REV_{it} - \Delta REC_{it})/TA_{it-1}] \\ & + \phi_3(PPE_{it}/TA_{it-1}) + \phi_4 Q1_{it} + \phi_5 Q2_{it} + \phi_6 Q3_{it} \\ & + \phi_7 Q4_{it} + \varepsilon_{it}, \end{aligned} \tag{1}$$

where TOT_ACC is total accruals (i.e., earnings before extraordinary items and discontinued operations less operating cash flows); TA is total assets; ΔREV is the quarterly change in revenues; ΔREC is the quarterly change in accounts receivable; PPE is gross property, plant, and equipment; and Q1, Q2, Q3, and Q4 are fiscal-quarter indicators. The (ΔREV − ΔREC) term controls for normal working capital accruals related to sales, the PPE term

TABLE 2
Descriptive Statistics for Selected Firm Characteristics and Rank Correlations with Earnings Surprise

Firm Characteristic	*Mean*	*Standard Deviation*	*First Quartile*	*Median*	*Third Quartile*	*Spearman Rank Correlation with SURPRISE*
Market value of common equity ($)	2,311.9	6,267.0	161.7	459.0	1,453.4	0.11 ***
Total assets ($)	1,811.8	4,397.9	129.1	371.1	1,251.2	0.06 ***
Sales ($)	470.9	1,028.1	37.5	111.9	356.6	0.08 ***
Traded on NASDAQ (%)	44.1	49.7	0.0	0.0	100.0	−0.03 ***
Ratio of fixed assets to total assets (%)	32.5	22.4	14.9	26.5	45.8	−0.01
Ratio of total debt to total assets (%)	61.7	111.9	5.2	38.2	85.9	−0.02 ***
Net income ($)	24.5	78.9	0.8	4.5	16.6	0.21 ***
Reporting losses (%)	18.6	38.9	0.0	0.0	0.0	−0.19 ***
Operating cash flows ($)	102.0	322.5	0.6	12.7	59.0	0.11 ***

*** denotes significant at the 0.01 level, based on two-tailed tests.

Tabulated amounts are based on data for quarter t. Dollar ($) amounts are in millions. The sample consists of 35,950 firm-quarters pertaining to 3,649 firms included in both the Compustat and I/B/E/S databases with complete data for our primary tests over 1993–1999, excluding utilities and financial services firms (two-digit SIC codes 49 and 60–67). SURPRISE is the I/B/E/S actual EPS for quarter t less the consensus forecast for quarter t, both rounded to the nearest penny. The consensus forecast is the mean of analysts' most recent EPS forecasts for quarter t available on I/B/E/S prior to the earnings announcement for quarter t. We combine SURPRISE ≤ −5¢ into one category and SURPRISE ≥ 5¢ into another. Fixed assets consist of property, plant, and equipment, net of accumulated depreciation.

controls for normal depreciation and related deferred tax accruals, the TA deflator controls for scale, and the fiscal-quarter indicators control for seasonality.

Equation (1) scales abnormal accruals by lagged TA, so we unscale them by multiplying the regression residuals by their corresponding lagged TA before accumulating them backward in time. To avoid inducing a spurious correlation between prior cumulative (unscaled) abnormal accruals and NOA, we rescale the cumulative abnormal accruals by sales for quarter $t - 1$, the same deflator in NOA.

Table 3 reports mean levels of (rescaled) abnormal accruals accumulated over the prior 4, 8, 12, 16, and 20 quarters, across quintiles of adjusted NOA. Consistent with the estimation of Equation (1), we adjust NOA only in this table by subtracting the mean NOA in the same two-digit SIC code and fiscal year. The table shows that abnormal accruals accumulated over up to 20 previous quarters are larger (at the 0.10 significance level or better) for firms in the upper quintile of adjusted NOA than for firms in the lower quintile. These results suggest that prior optimistic accounting choices are associated with larger levels of net operating assets relative to sales.

Model

We consider two issues in developing a model to test H_a. First, recall that we define SURPRISE as actual EPS less the consensus EPS forecast. Because we measure SURPRISE in pennies, treating it as an ordinal variable, we use an estimation model specifically designed for ordinal dependent variables. Second, we suspect that the effects of the independent variables vary across various earnings surprise benchmarks; for example, managers may have less accounting discretion or weaker incentives to beat expectations by 5¢ or more than to beat expectations by just 1¢. Our model thus allows the parameters of the independent variables to vary across predetermined earnings surprise benchmarks. We incorporate both of these design choices using the following generalized ordered logit model, derived in the Appendix:

$$\Pr(\mathrm{SURPRISE}_{it} \geq k)/\Pr(\mathrm{SURPRISE}_{it} < k) = \exp(\beta_{0,k} + \beta_{1,k}\mathrm{NOA}_{it} + \boldsymbol{\beta}_{\mathbf{k}}'\mathbf{CONTROLS}_{it}) \qquad (2)$$

where the left-hand-side expression is the odds of reporting an earnings surprise of at least k¢, a predetermined benchmark; SURPRISE is the signed EPS surprise; NOA is our proxy for overstated net assets already on the balance sheet; **CONTROLS** is a vector of control variables; βs are parameters allowed to vary with k; and i and t denote firm and quarter. For m categories in SURPRISE, Equation (2) yields $m - 1$ uniquely identified equations that can be estimated jointly through maximum likelihood techniques.[5] Thus, because we coded SURPRISE using 11 categories (i.e., $-5, -4, -3, \ldots, 3, 4, 5$), our empirical implementation of Equation (2) will yield a set of parameter estimates for each of ten uniquely identified equations.

H_a predicts a negative association between SURPRISE and NOA. Therefore, we expect the coefficients on NOA to be negative in all ten equations. For example, consider $k = 0$, a zero earnings surprise. For $k = 0$, a negative coefficient $\beta_{1,0}$ on NOA implies that higher values of NOA are associated with lower odds of a zero or positive earnings surprise. Equation (2) also allows us to estimate how changes in the independent variables affect the

---

[5] For further details on generalized ordered logit models and their estimation, see Clogg and Shihadeh (1994, 146–147) and Fu (1998).

TABLE 3
Average Prior Cumulative Abnormal Accruals across Quintiles of Net Operating Assets (Scaled by Sales)

Accumulation Period (Quarters)	*n*	*Quintiles of NOA*					*Difference between Upper and Lower Quantiles*	*Difference t-statistic*
		Upper (1)	*(2)*	*(3)*	*(4)*	*Lower (5)*		
[t − 1, t − 4]	26,146	0.32	−0.07	−0.12	−0.14	−0.06	0.38	11.65 ***
[t − 1, t − 8]	21,104	0.19	−0.22	−0.27	−0.28	−0.22	0.41	8.29 ***
[t − 1, t − 12]	17,478	0.08	−0.38	−0.41	−0.40	−0.34	0.42	6.44 ***
[t − 1, t − 16]	14,969	−0.17	−0.50	−0.52	−0.50	−0.46	0.29	3.81 ***
[t − 1, t − 20]	11,555	−0.43	−0.63	−0.59	−0.57	−0.58	0.15	1.63 *

*** and * denote significant at the 0.01 and 0.10 levels, respectively, based on one-tailed tests. The t-statistics for differences in means between upper and lower quintiles are based on unequal variances across quintiles.

The initial sample consists of 35,950 firm-quarters pertaining to 3,649 firms included in both the Compustat and I/B/E/S databases with complete data for our primary tests over 1993–1999, excluding utilities and financial services firms (two-digit SIC codes 49 and 60–67). Of these firm-quarters, only n have data to estimate cumulative abnormal accruals over a particular accumulation period. NOA is net operating assets (i.e., shareholders' equity less cash and marketable securities, plus total debt) at the beginning of quarter t, scaled by sales for quarter t − 1. Consistent with the estimation of Equation (1) below, we adjust NOA only in this table by subtracting the mean NOA in the same two-digit SIC code and fiscal year. We estimate abnormal accruals for firm i in quarter t using the residual from the following regression, estimated by two-digit SIC code and fiscal year using 1988–1999 data (see Jones 1991; Dechow et al. 1995; Han and Wang 1998):

(*Continued on next page*)

TABLE 3 (Continued)

$$TOT\_ACC_{it}/TA_{it-1} = \phi_1(1/TA_{it-1}) + \phi_2[(\Delta REV_{it} - \Delta REC_{it})/TA_{it-1}] + \phi_3(PPE_{it}/TA_{it-1}) + \phi_4 Q1_{it} + \phi_5 Q2_{it} + \phi_6 Q3_{it} + \phi_7 Q4_{it} + \varepsilon_{it}, \quad (1)$$

where:

TOT_ACC = total accruals (i.e., earnings before extraordinary items and discontinued operations less operating cash flows);
TA = total assets;
ΔREV = quarterly change in revenues;
ΔREC = quarterly change in accounts receivable;
PPE = gross property, plant, and equipment; and
Q1, Q2, Q3, and Q4 = fiscal-quarter indicators.

The abnormal accruals from Equation (1) are implicitly scaled by lagged TA, so we unscale them by multiplying the regression residuals by the corresponding lagged TA before accumulating them backward in time. To avoid inducing a spurious correlation between prior cumulative abnormal accruals and NOA, we rescale the cumulative abnormal accruals by sales for quarter $t - 1$, the same deflator in NOA. The tabulated amounts are mean levels of (rescaled) abnormal accruals accumulated over the prior 4, 8, 12, 16, and 20 quarters, across quintiles of adjusted NOA.

odds of a given earnings surprise. For instance, an increase of δ in NOA changes the odds of meeting or beating expectations by $100[\exp(\delta \times \beta_{1,0}) - 1]$ percent. We derive the expression for the percentage change in odds in the Appendix.[6]

The vector **CONTROLS** includes variables capturing other constraints on earnings management, managerial incentives to meet or slightly beat forecasts, and firm performance and size. We discuss these in turn.

One constraint on earnings management is the number of shares outstanding. Managers of firms with more shares outstanding may find it more difficult to manage earnings toward expectations, because a penny short in EPS translates into more dollars of actual earnings for a firm with more shares outstanding than for a firm with fewer shares outstanding. To capture this effect, we include the weighted average number of common shares outstanding during quarter t (SHARES), the denominator of basic EPS, and expect its coefficients to be negative.[7]

Another constraint on earnings management is audit quality. Firms with Big 5 auditors have lower levels of abnormal accruals (Becker et al. 1998; Payne and Robb 2000). However, Libby and Kinney (2000) find that few Big 5 auditors expect a client to correct otherwise quantitatively immaterial income-increasing misstatements if such corrections would result in the client missing earnings forecasts. Therefore, we include an indicator variable coded 1 (0 otherwise) if the firm has a Big 5 auditor (BIG5) in quarter t, but make no prediction about the sign of its coefficients.

Managers are more likely to report earnings that meet or beat expectations if they have strong incentives to do so—when their firms have a high price-to-book ratio (Skinner and Sloan 2001), high litigation risk (Soffer et al. 2000; Matsumoto 2002), a large analyst following (Johnson 1999), or a previous pattern of meeting or beating expectations.[8] We measure the price-to-book ratio (PB) as the market value of common equity divided by the book value of shareholders' equity, both at the end of quarter t. We measure litigation risk (LTGN_RISK) as an indicator variable coded 1 (0 otherwise) if the firm is in one of four industries identified in prior research (e.g., Francis et al. 1994; Ali and Kallapur 2001) as most susceptible to securities litigation—pharmaceuticals/biotechnology (SIC codes 2833–2836, 8731–8734), computers (3570–3577, 7370–7374), electronics (3600–3674), and retail (5200–5961). We measure analyst following (ANALYSTS) as the number of analysts in the I/B/E/S consensus EPS forecast for quarter t. Last, we measure the previous earnings surprise pattern (PREV_MB) as an indicator variable coded 1 (0 otherwise) if, based on I/B/E/S, the firm reported a nonnegative earnings surprise (i.e., SURPRISE ≥ 0) in quarter $t - 1$. We expect the coefficients on these variables to be positive.

Managers are more likely to report earnings that miss expectations if the expectations are imprecise (Payne and Robb 2000). We measure this imprecision by the coefficient of variation in analysts' most recent forecasts for quarter t (CV_FORECAST), and expect the coefficients on this variable to be negative.[9]

---

[6] The advantage of interpreting each variable's effect in terms of a percentage change in odds rather than a discrete or marginal change in probabilities is that the change in odds for that variable does not depend on the levels of other variables in the model.

[7] Alternatively, managers may meet or beat expectations by reducing the denominator of EPS through share repurchases. We do not distinguish between these two reasons since they both predict negative coefficients on SHARES.

[8] Including the firm's price-to-earnings ratio instead of its price-to-book ratio yields similar results.

In addition to managing accruals, managers may report earnings that meet or beat expectations by "talking down" analysts' forecasts (Matsumoto 2002), that is, by guiding analysts' expectations down prior to announcing earnings. If this strategy is successful, then we expect firms with downward forecast revisions to be more likely to at least meet expectations. However, if analysts underreact to the "bad" news (e.g., Abarbanell and Bernard 1992; Easterwood and Nutt 1999), then firms with downward forecast revisions may still miss expectations. We measure downward forecast revisions (DOWN_REV) as an indicator variable coded 1 (0 otherwise) if at least one analyst revised his or her forecast down after the earnings announcement for quarter $t - 1$ but before the earnings announcement for quarter t (Bartov et al. 2002). We make no prediction for the sign of the coefficients on this variable.

Previous research (e.g., Abarbanell and Lehavy 2001a; Skinner and Sloan 2001) suggests that the level of earnings surprise is increasing in firm performance. We control for performance by including sales growth (SALES_GRW), defined as sales for quarter t divided by sales for quarter $t - 4$, less 1; return on equity (ROE), defined as net income for quarter t divided by shareholders' equity at the end of quarter t; and the change in the return on equity (ΔROE), defined as ROE in quarter t less ROE in quarter $t - 4$. We expect the coefficients on these variables to be positive.

Finally, we control for firm size because analysts tend to issue less optimistic forecasts for larger firms (Das et al. 1998). We measure firm size (MKT_CAP) as the natural logarithm of the market value of common equity at the end of quarter t, and expect the coefficients on this variable to be positive.

IV. EMPIRICAL ANALYSES

Descriptive Statistics and Correlations

The first five columns of Table 4 report descriptive statistics for the independent variables. The mean (median) level of NOA is 2.66 (1.97), suggesting that net operating assets are about twice as large or larger as sales for most firm-quarters. Sample firms have on average 61 million shares outstanding, and 97 percent of them have a Big 5 auditor. The mean price-to-book ratio is 3.42; about 33 percent of the firms are in highly litigious industries; the average number of analysts following a firm is 6.39; and firms met or beat analysts' expectations in the previous quarter 68 percent of the time. The mean (median) coefficient of variation in analysts' forecasts is 0.20 (0.06), indicating a high degree of consensus among analysts in most firm-quarters. In 23 percent of the cases, at least one analyst revised downward his or her forecast for current-quarter EPS. Average sales growth is 21 percent; however, the mean ROE is 8 percent, about 2 percent lower than ROE for the same quarter in the previous year. The mean MKT_CAP is 6.28, a market value of common equity of about $534 million.

The last column of Table 4 reports Spearman rank correlations between SURPRISE and each independent variable. As predicted, the correlation between SURPRISE and NOA is negative ($r = -0.11$, one-tailed $p < 0.01$). The correlations between SURPRISE and the

---

[9] Higher values of CV_FORECAST also are consistent with less accurate analysts' forecasts. Analysts tend to issue less accurate and more optimistically biased forecasts for firms with more volatile earnings (Das et al. 1998; Lim 2001). As a sensitivity check, we include a control variable for earnings volatility, measured as the coefficient of variation in I/B/E/S actual EPS over the previous eight quarters. The coefficients on this variable are significantly negative at the 0.10 level or better; the coefficients on the other variables remain qualitatively unchanged.

TABLE 4
Descriptive Statistics for Independent Variables and Rank Correlations with Earnings Surprise

Independent Variable	*Mean*	*Standard Deviation*	*First Quartile*	*Median*	*Third Quartile*	*Predicted Sign*	*Spearman Rank Correlation with SURPRISE*
NOA	2.66	2.63	1.26	1.97	2.99	–	−0.11 ***
SHARES	60.69	116.70	12.03	23.50	52.40	–	−0.04 ***
BIG5	0.97	0.18	1.00	1.00	1.00	?	0.01 *
PB	3.42	3.57	1.60	2.49	3.99	+	0.12 ***
LTGN_RISK	0.33	0.47	0.00	0.00	1.00	–	−0.01 ***
ANALYSTS	6.39	4.71	3.00	5.00	8.00	+	0.06 ***
PREV_MB	0.68	0.47	0.00	1.00	1.00	+	0.25 ***
CV_FORECAST	0.20	0.43	0.03	0.06	0.17	–	−0.13 ***
DOWN_REV	0.23	0.42	0.00	0.00	0.00	?	−0.17 ***
SALES_GRW	0.21	0.42	0.02	0.12	0.30	+	0.18 ***
ROE	0.08	0.27	0.04	0.12	0.19	+	0.13 ***
ΔROE	−0.02	0.30	−0.07	−0.01	0.04	+	0.17 ***
MKT_CAP	6.28	1.60	5.09	6.13	7.28	+	0.11 ***

(Continued on next page)

TABLE 4 (Continued)

*** and * denote significant at the 0.01 and 0.10 levels, respectively, based on one-tailed tests for signed predictions and two-tailed tests otherwise.

The sample consists of 35,950 firm-quarters pertaining to 3,649 firms included in both the Compustat and I/B/E/S databases with complete data over 1993–1999, excluding utilities and financial services firms (two-digit SIC codes 49 and 60–67). The variables are defined as follows:

SURPRISE = I/B/E/S actual EPS for quarter t less the consensus forecast for quarter t, both rounded to the nearest penny. The consensus forecast is the mean of analysts' most recent EPS forecasts for quarter t available on I/B/E/S prior to the earnings announcement for quarter t. We combine SURPRISE ≤ -5¢ into one category and SURPRISE ≥ 5¢ into another;

NOA = net operating assets (i.e., shareholders' equity less cash and marketable securities, plus total debt) at the beginning of quarter t, scaled by sales for quarter t − 1;

SHARES = weighted average number of common shares outstanding during quarter t;

BIG5 = indicator variable coded 1 if the firm has a Big 5 auditor in quarter t, 0 otherwise;

PB = market value of common shares divided by shareholders' equity, both at the end of quarter t;

LTGN_RISK = indicator variable coded 1 if the firm is in one of the following industries: pharmaceuticals/biotechnology (SIC codes 2833–2836, 8731–8734), computers (3570–3577, 7370–7374), electronics (3600–3674), or retail (5200–5961), 0 otherwise;

ANALYSTS = number of analysts in the I/B/E/S consensus EPS forecast for quarter t;

PREV_MB = indicator variable coded 1 if, based on I/B/E/S, the firm reported a nonnegative earnings surprise (i.e., SURPRISE ≥ 0) in quarter t − 1, 0 otherwise;

CV_FORECAST = coefficient of variation in analysts' most recent forecasts for quarter t;

DOWN_REV = indicator variable coded 1 if at least one of the firm's analysts revised his or her forecast down prior to the end of quarter t but after the earnings announcement date for quarter t − 1, 0 otherwise;

SALES_GRW = sales for quarter t divided by sales for quarter t − 4, less 1;

ROE = net income for quarter t divided by shareholders' equity at the end of quarter t;

ΔROE = ROE for quarter t less ROE for quarter t − 1; and

MKT_CAP = natural logarithm of market value of common shares at the end of quarter t.

remaining variables are significant at conventional levels. Untabulated analyses reveal that rank correlations between most independent variables also are significant at conventional levels.

Regression Results

Table 5 reports estimations results for Equation (2). Recall that the model allows the effects of the independent variables to vary across levels of predetermined earnings surprise benchmarks, k. Because SURPRISE consists of 11 categories coded -5 through 5, we obtain ten sets of uniquely identified parameter estimates, one for each k ranging from -4 through 5.[10] We estimate the model using the generalized ordered logit regression technique described in Fu (1998). The test statistics for all coefficients are heteroscedasticity-consistent and, because our sample includes multiple observations for most firms, they are adjusted for residual correlation among observations for the same firm.

Panel A of Table 5 presents complete estimation results when the dependent variable is the odds of at least meeting earnings expectations, that is, when $k = 0$. For each independent variable, the last column of Panel A presents results of Wald tests proposed by Brant (1990) for the null hypothesis that the coefficients on the independent variable are constant across levels of k. These tests reject the null at the 0.01 level for all variables except BIG5. Therefore, we also report summary results for all values of k. Specifically, Panel B shows the percentage changes in odds for each variable across k.[11] For a continuous variable, the change in odds is based on a standard-deviation increase in the variable; for indicators, it is based on a change from 0 to 1 (see the Appendix).

Panel A of Table 5 shows that the coefficient on NOA is negative ($p < 0.01$). A standard-deviation increase in the beginning balance of net operating assets relative to sales decreases by 7.7 percent the odds of at least meeting the consensus forecast, all else equal. That is, consistent with H_a, the likelihood of meeting or beating analysts' earnings forecasts by optimistically biasing earnings decreases with the extent to which net assets are already overstated on the balance sheet.

With respect to the control variables, the coefficient on SHARES is negative ($p < 0.01$), suggesting that the odds of at least meeting expectations decrease with shares outstanding. The insignificant coefficient on BIG5 suggests that our proxy for audit quality is unrelated to the odds of at least meeting analysts' expectations. The coefficients on PB, LTGN_RISK, ANALYSTS, and PREV_MB are positive and, except for LTGN_RISK, significant at the 0.10 level or better. These results suggest that the odds of at least meeting expectations are increasing in the firm's price-to-book ratio, its record of meeting or beating forecasts in the previous quarter, and the size of its analyst following. The coefficients on CV_FORECAST and DOWN_REV are negative ($p < 0.01$), suggesting that the odds of at least meeting expectations decrease with the dispersion in analysts' forecasts and with the presence of downward forecast revisions. Finally, the coefficients on SALES_GRW, ROE, ΔROE, and MKT_CAP are positive ($p < 0.01$), suggesting that the odds of at least meeting expectations increase with the firm's performance and size.

Panel B of Table 5 shows that the percentage changes in odds for NOA are negative ($p < 0.01$) across all k. A standard-deviation increase in the beginning balance of net

---

[10] The odds of SURPRISE ≥ -5¢ can be recovered from the odds of SURPRISE ≥ -4¢ through tedious algebra. Interested readers may contact us.

[11] We report percentage changes in odds because they capture both the sign and relative effect of each independent variable. Therefore, unlike coefficients, percentage changes in odds are comparable across independent variables.

TABLE 5
Regression Results for Generalized Ordered Logit Model

Model:

$$Pr(SURPRISE_{it} \geq k)/Pr(SURPRISE_{it} < k) = exp(\beta_{0,k} + \beta_{1,k}NOA_{it} + \beta_{2,k}SHARES_{it} + \beta_{3,k}BIG5_{it} + \beta_{4,k}PB_{it} + \beta_{5,k}LTGN\_RISK_{it}$$
$$+ \beta_{6,k}ANALYSTS_{it} + \beta_{7,k}PREV\_MB_{it} + \beta_{8,k}CV\_FORECAST_{it} + \beta_{9,k}DOWN\_REV_{it}$$
$$+ \beta_{10,k}SALES\_GRW_{it} + \beta_{11,k}ROE_{it} + \beta_{12,k}\Delta ROE_{it}$$
$$+ \beta_{13,k}MKT\_CAP_{it} + v_{it})$$

Panel A: Regression Results for k = 0¢, i.e., Odds of Meeting or Beating vs. Missing Analysts' Forecasts Pr(SURPRISE ≥ 0)/Pr(SURPRISE < 0)

Independent Variable	*Predicted Sign*	*Coefficient*	*z-statistic*	*Change in Odds (%)*[a]	*Brant* $\chi^2_{(9\ df)}$[b]
Intercept	?	−0.805	−7.25***	NA	NA
NOA	−	−0.031	−4.98***	−7.7	30.73***
SHARES	−	−0.001	−5.60***	−11.3	47.37***
BIG5	?	0.013	0.15	1.3	12.68
PB	+	0.026	4.86***	9.7	289.65***
LTGN_RISK	+	0.030	0.94	3.0	198.16***
ANALYSTS	+	0.008	1.62*	3.6	95.44***
PREV_MB	+	1.004	33.18***	172.9	554.33***
CV_FORECAST	−	−0.244	−7.92***	−9.9	1,580.36***
DOWN_REV	?	−0.594	−19.82***	−44.8	229.10***
SALES_GRW	+	0.606	13.63***	28.9	92.13***
ROE	+	0.228	2.87***	6.4	493.25***
ΔROE	+	0.156	2.40***	4.7	77.42***

(Continued on next page)

TABLE 5 (Continued)

Independent Variable	*Predicted Sign*	*Coefficient*	*z-statistic*	*Change in Odds (%)*[a]	*Brant* $\chi^2_{(9\ df)}$[b]
MKT_CAP	+	0.156	10.17***	28.4	42.05***
Likelihood ratio $\chi^2_{(130\ df)}$		8,105.61***			
Maximum likelihood R^2		0.20			

Panel B: Percentage Change in Odds of Reporting SURPRISE ≥ k, Given a Change in the Independent Variable[a]

Independent Variable	*Pred. Sign*	*k*									
		−4¢	*−3¢*	*−2¢*	*−1¢*	*0¢*	*1¢*	*2¢*	*3¢*	*4¢*	*5¢*
NOA	−	−6.2***	−6.8***	−7.0***	−7.9***	−7.7***	−7.9***	−7.5***	−7.0***	−5.7***	−6.6***
SHARES	−	−0.6	−1.1	−1.1	−6.2**	−11.3***	−13.1***	−14.2***	−17.3***	−18.1***	−20.3***
BIG5	?	−6.9	−8.7	−2.3	−6.3	1.3	9.8	2.3	4.2	7.1	5.0
PB	+	24.3***	21.8***	17.7***	15.0***	9.7***	−3.2	−8.6•	−12.1•	−13.6•	−14.5•
LTGN_RISK	+	4.1	3.8	4.2	3.8	3.0	−14.9•	−20.4•	−27.6•	−32.6•	−36.5•
ANALYSTS	+	8.1**	9.1***	10.0***	8.1***	3.6*	−4.0	−6.0•	−6.1•	−5.3•	−8.0•
PREV_MB	+	186.8***	186.1***	186.7***	188.5***	172.9***	80.6***	54.4***	41.8***	37.1***	35.7***
CV_FORECAST	−	−19.7***	−19.0***	−17.2***	−14.2***	−9.9***	−4.1***	1.3	5.1•	8.0•	10.0•
DOWN_REV	?	−40.3***	−42.0***	−45.9***	−47.3***	−44.8***	−33.3***	−31.7***	−31.2***	−29.6***	−25.9***
SALES_GRW	+	27.8***	27.7***	28.9***	28.0***	28.9***	21.3***	16.4***	15.1***	12.7***	11.6***
ROE	+	23.9***	21.4***	16.3***	11.9***	6.4***	−3.2	−9.4•	−14.0•	−17.9•	−21.3•
ΔROE	+	1.8	1.9	2.2	3.0*	4.7***	12.3***	15.1***	16.2***	18.1***	19.1***
MKT_CAP	+	15.5***	14.8***	16.3***	22.1***	28.4***	22.0***	23.7***	26.5***	28.4***	34.4***

(*Continued on next page*)

TABLE 5 (Continued)

***,**, and * denote significant at the 0.01, 0.05, and 0.10 levels, respectively, based on one-tailed test for signed predictions and two-tailed tests otherwise. The z-statistics are heteroscedasticity-consistent and adjusted for residual correlation among observations pertaining to the same firm.

• denotes significant at the 0.10 level or better, based on two-tailed tests, but opposite to our predicted sign.

[a] The percentage change in odds is the effect of a change in the independent variable on the odds of reporting SURPRISE $\geq$ k instead of SURPRISE $<$ k, where k is a predetermined earnings surprise benchmark. For continuous variables, the percentage change in odds is $100[\exp(s_j\beta_{j,k}) - 1]$, where s_j is the sample standard deviation of variable j (reported in Table 4) and $\beta_{j,k}$ is the estimated regression coefficient for variable j with respect to earnings surprise benchmark k. For indicator variables, the percentage change in odds is $100[\exp(\beta_{j,k}) - 1]$.

[b] The Brant χ^2 statistic tests the null hypothesis that odds for the independent variable are proportional, i.e., that the coefficients on the independent variable are constant across k (Brant 1990).

The sample consists of 35,950 firm-quarters pertaining to 3,649 firms included in both the Compustat and I/B/E/S databases with complete data for our primary tests over 1993–1999, excluding utilities and financial services firms (two-digit SIC codes 49 and 60–67).

The variables are defined in Table 4.

operating assets relative to sales decreases the odds of reporting a larger positive or smaller negative earnings surprise by 5.7 to 7.9 percent. These results further support H_a.

Panel B of Table 5 also shows that some of the independent variables have a consistent pattern across all k. For example, the percentage changes in odds for BIG5 are not significant at conventional levels for any value of k, suggesting that our proxy for audit quality is unrelated to earnings surprises. The percentage changes in odds for PREV_MB, SALES_GRW, and MKT_CAP are positive ($p < 0.01$) across all values of k, suggesting that the odds of reporting larger positive or smaller negative earnings surprises are increasing in the likelihood of a nonnegative earnings surprise in the previous quarter, sales growth, and firm size. The percentage changes in odds for DOWN_REV are negative ($p < 0.01$) across all k, suggesting that the odds of reporting larger positive or smaller negative earnings surprises decrease when analysts make downward forecast revisions.

Other independent variables have a consistent pattern across only part of the range of values for k. For example, the percentage changes in odds for ΔROE are positive ($p < 0.10$) when $k \geq -1$, suggesting that the odds of reporting larger positive surprises are increasing in ΔROE, but the odds of reporting earnings surprises lower than −1¢ are unrelated to ΔROE. The percentage changes in odds for SHARES are negative ($p < 0.05$) when $k \geq -1$, suggesting that more shares outstanding decrease the odds of reporting larger positive surprises, but have no effect on the odds of missing expectations by more than 1¢. The percentage changes in odds for LTGN_RISK are negative ($p < 0.10$) when $k \geq 1$, suggesting that litigation risk decreases the odds of larger positive surprises, but has no effect on the odds of missing expectations.

Finally, the pattern of some variables switches signs around positive earnings surprises of 1¢ or 2¢. For example, the percentage changes in odds for PB, ANALYSTS, and ROE are positive ($p < 0.10$) when $k < 1$ but negative ($p < 0.10$) when $k > 1$. These results suggest that firms with high price-to-book ratios, large analyst following, and high returns on equity are more likely to miss expectations by smaller amounts than by larger amounts, and they are less likely to beat expectations by larger amounts than by smaller amounts. In contrast, the percentage changes in odds for CV_FORECAST are negative ($p < 0.01$) when $k < 2$ but positive ($p < 0.10$) when $k > 2$, suggesting that firms with imprecise forecasts are more likely to miss expectations by larger than smaller amounts, and are more likely to beat expectations by more than 2¢.

In sum, the results reported in Table 5 are consistent with H_a, showing a negative association between earnings surprises and our proxy for overstated net asset values. We interpret these results as evidence that the likelihood of reporting larger positive or smaller negative earnings surprises decreases with the extent to which net assets are already overstated on the balance sheet. That is, prior optimism in financial reporting accumulates on the balance sheet, and reduces managers' ability to optimistically bias earnings in future periods, if managers want to stay within GAAP and its implementation guidance. The results also indicate that the relationship between earnings surprises and the independent variables is complex. The presence of a Big 5 auditor is consistently unrelated to earnings surprises. However, none of the remaining variables measuring constraints on earnings management, managerial incentives to meet or beat analysts' forecasts, firm performance, and size has a constant effect across earnings surprise levels.

Sensitivity Analyses

Alternative Econometric Techniques

We use a generalized ordered logit regression of SURPRISE on NOA and **CONTROLS** because SURPRISE is an ordinal dependent variable and because we allow the coefficients on all independent variables to vary across levels of SURPRISE. Using ordinary least squares (OLS) or "regular" ordered logit regressions constrains the coefficients on each variable to be constant across levels of SURPRISE, yielding only one set of coefficient estimates. Moreover, OLS regression treats SURPRISE as a continuous dependent variable.

To assess the robustness of our results under alternative econometric specifications that constrain the coefficients on each variable to be constant across levels of SURPRISE, we regress SURPRISE on NOA and **CONTROLS** using OLS and "regular" ordered logit regressions. The coefficient on NOA is -0.053 ($t = -8.54$, $p < 0.01$) based on OLS regression; it is -0.028 ($z = -4.31$, $p < 0.01$) based on "regular" ordered logit regression. Thus, the results bearing on our hypothesis are robust to these alternative specifications. However, OLS and "regular" ordered logit regressions lead to substantial changes in the coefficients on some of the other independent variables.

Composition of Net Operating Assets

To capture the possibility that different components of net operating assets are subject to varying degrees of managerial manipulation (Beatty et al. 1995; Hunt et al. 1996; Teoh et al. 1998; Gaver and Paterson 1999), we split NOA into three components:

1) Working capital (WC), defined as current assets less cash, marketable securities, and current liabilities, plus short-term debt and the current portion of long-term debt, all at the beginning of quarter t and scaled by sales for quarter $t - 1$.
2) Net fixed assets (FIXED_ASSETS), defined as property, plant, and equipment, net of accumulated depreciation, at the beginning of quarter t and scaled by sales for quarter $t - 1$.
3) Other long-term assets (OTHER_LTASSETS), defined as NOA less WC and FIXED_ASSETS.

We separate working capital from long-term net operating assets because prior research suggests that current accruals are likely easier to manage than long-term accruals (Hunt et al. 1996; Beneish 1998; Teoh et al. 1998). We separate fixed assets from other long-term assets because managing earnings through depreciation is transparent (e.g., firms are required to disclose the effects of changes in depreciation policies) or costly if managers must time capital investments to make depreciation-related earnings management less transparent (Beneish 1998).

Table 6 reports selected regression results using WC, FIXED_ASSETS, and OTHER_LTASSETS in place of NOA in Equation (2). Consistent with H_a and the results reported in Table 5, the coefficients on the three components of net operating assets are negative and significant at the 0.05 level or better across the various surprise benchmarks k. Brant (1990) χ^2 tests reject at the 0.01 level the null hypothesis that the coefficients on each variable are constant across k. Finally, the coefficients on WC are about 9 to 28 (3 to 5) times larger in absolute terms than the coefficients on FIXED_ASSETS (OTHER_LTASSETS), suggesting that, relative to the level of long-term net assets, the level of working capital has a stronger effect on the odds of reporting a predetermined earnings surprise.

Additional Robustness Tests

The level of net operating assets relative to sales tends to vary across industries (Nissim and Penman 2001). To control for this effect, we reestimate Equation (2) including indicators for two-digit SIC codes. The coefficients on NOA (not tabulated) remain negative and significant at the 0.01 level.

For reporting quarterly performance, GAAP requires managers to estimate annual operating expenses and allocate these costs across quarters (e.g., APB 1973; FASB 1974). This opportunity to misreport quarterly performance is exacerbated by the fact that financial

P-10,932

TABLE 6
Selected Regression Results for Generalized Ordered Logit Model Using Components of Net Operating Assets

Model:

$$Pr(SURPRISE_{it} \geq k)/Pr(SURPRISE_{it} < k) = exp(\gamma_{0,k} + \gamma_{1,k}WC_{it} + \gamma_{2,k}FIXED\_ASSETS_{it} + \gamma_{3,k}OTHER\_LTASSETS_{it} + \gamma_{4,k}SHARES_{it}$$
$$+ \gamma_{5,k}BIG5_{it} + \gamma_{6,k}PB_{it} + \gamma_{7,k}LTGN\_RISK_{it} + \gamma_{8,k}ANALYSTS_{it} + \gamma_{9,k}PREV\_MB_{it}$$
$$+ \gamma_{10,k}CV\_FORECAST_{it} + \gamma_{11,k}DOWN\_REV_{it} + \gamma_{12,k}SALES\_GRW_{it} + \gamma_{13,k}ROE_{it}$$
$$+ \gamma_{14,k}\Delta ROE_{it} + \gamma_{15,k}MKT\_CAP_{it} + \nu_{it})$$

		WC			*FIXED_ASSETS*			*OTHER_LTASSETS*		
k	*Predicted Sign*	*Coefficient*	*z-statistic*	*Change in Odds (%)*[a]	*Coefficient*	*z-statistic*	*Change in Odds (%)*[a]	*Coefficient*	*z-statistic*	*Change in Odds (%)*[a]
−4¢	−	−0.204	−6.28***	−12.0	−0.018	−2.22**	−4.2	−0.058	−3.26***	−6.4
−3	−	−0.223	−7.39***	−13.0	−0.020	−2.67***	−4.8	−0.060	−3.55***	−6.5
−2	−	−0.226	−8.04***	−13.2	−0.021	−3.04***	−5.0	−0.063	−4.08***	−6.9
−1	−	−0.245	−9.26***	−14.2	−0.026	−4.04***	−6.2	−0.060	−4.40***	−6.6
0	−	−0.263	−10.56***	−15.1	−0.024	−3.81***	−5.8	−0.049	−3.72***	−5.4
1	−	−0.296	−10.49***	−16.9	−0.021	−3.20***	−5.0	−0.103	−7.57***	−11.0
2	−	−0.370	−10.83***	−20.6	−0.020	−2.74***	−4.7	−0.133	−7.58***	−14.0
3	−	−0.420	−10.72***	−23.1	−0.020	−2.51***	−4.7	−0.147	−6.94***	−15.3
4	−	−0.451	−9.92***	−24.6	−0.016	−1.96**	−3.9	−0.142	−6.02***	−14.8
5	−	−0.484	−10.07***	−26.1	−0.018	−2.09**	−4.4	−0.155	−5.81***	−16.1
Brant $\chi^2_{(9\ df)}$[b]			92.81***			108.03***			84.34***	

(*Continued on next page*)

TABLE 6 (Continued)

*** and** denote significant at the 0.01 and 0.05 levels, respectively, based on one-tailed test for signed predictions and two-tailed tests otherwise. The z-statistics are heteroscedasticity-consistent and adjusted for residual correlation among observations pertaining to the same firm.

[a] The percentage change in odds is the effect of a change in the independent variable on the odds of reporting SURPRISE $\geq$ k instead of SURPRISE $<$ k, where k is a predetermined earnings surprise benchmark. For continuous variables, the percentage change in odds is $100[\exp(s_j\beta_{j,k}) - 1]$, where s_j is the sample standard deviation of variable j (reported in Table 4) and $\beta_{j,k}$ is the estimated regression coefficient for variable j with respect to earnings surprise benchmark k. For indicator variables, the percentage change in odds is $100[\exp(\beta_{j,k}) - 1]$.

[b] The Brant χ^2 statistic tests the null hypothesis that odds for the independent variable are proportional, i.e., that the coefficients on the independent variable are constant across k (Brant 1990).

The sample consists of 35,950 firm-quarters pertaining to 3,649 firms included in both the Compustat and I/B/E/S databases with complete data for our primary tests over 1993–1999, excluding utilities and financial services firms (two-digit SIC codes 49 and 60–67). The variables are defined as follows:

NOA = net operating assets (i.e., shareholders' equity less cash and marketable securities, plus total debt) at the beginning of quarter t, scaled by sales for quarter t − 1;
WC = current assets less cash, marketable securities, and current liabilities, plus current debt, all at the beginning of quarter t, scaled by sales for quarter t − 1;
FIXED_ASSETS = property, plant, and equipment, net of accumulated depreciation, at the beginning of quarter t, scaled by sales for quarter t − 1; and
OTHER_LTASSETS = NOA less WC less FIXED_ASSETS.

All other variables are defined in Table 4.

statements for the first three quarters are usually audited retrospectively during the annual external audit (Elliott and Shaw 1988; Ettredge et al. 2000). Consequently, security markets tend to react more strongly to negative earnings surprises in earlier quarters than in the fourth quarter (Mendenhall and Nichols 1988), suggesting that managers may have stronger incentives to meet or beat expectations earlier in the fiscal year. To capture this potential effect, we reestimate Equation (2) including fiscal-quarter indicators. The coefficients on NOA (not tabulated) remain negative and significant at the 0.01 level.

Because regulators and the media routinely mention the consensus analyst forecast as an earnings target, we calculate SURPRISE using the consensus forecast as a proxy for the market's earnings expectations. However, market expectations may be based on the most recent forecast (Brown and Kim 1991; Brown 2001). We recalculate SURPRISE using the most recent I/B/E/S forecast prior to the earnings announcement and reestimate Equation (2). Results (not tabulated) are similar to those reported in Table 5.

Abarbanell and Lehavy (2001b) point out that database providers such as I/B/E/S tend to adjust both actual and expected earnings for nonrecurring items. Because SURPRISE is based on I/B/E/S data and some of our independent variables (such as NOA) are based on Compustat data, we repeat the tests reported in Table 5 using (1) only firms for which I/B/E/S actual EPS agrees with Compustat EPS, and (2) only firms without special items on Compustat. The resulting coefficients on NOA (not tabulated) remain negative and significant at the 0.01 level.

Finally, prior research suggests that earnings management behavior and bias in analysts' forecasts differ between firms reporting losses and firms reporting profits (e.g., Degeorge et al. 1999; Easterwood and Nutt 1999; Brown 2001). We repeat the tests in Table 5 excluding firms reporting losses (1) in quarter t, (2) in quarter t − 1, and (3) in both quarters t and t − 1. For each set of tests, the coefficients on NOA (not tabulated) remain negative and significant at the 0.01 level.

V. CONCLUSION

Because the balance sheet accumulates the effects of previous accounting choices, the level of net operating assets partly reflects the extent of previous earnings management. We present evidence consistent with the hypothesis that managers' ability to optimistically bias earnings decreases with the extent to which net asset values are already overstated on the balance sheet.

Using 1993–1999 quarterly data for a sample of 3,649 nonfinancial, nonregulated firms, we estimate a generalized ordered logit model relating the level of a firm's earnings surprise to our proxy for the degree of overstatement in net asset values, other constraints on earnings management, managerial incentives to meet or beat analysts' earnings forecasts, and firm performance and size. We measure the degree of overstatement in net asset values using the beginning balance of net operating assets relative to sales, a proxy that additional analysis shows is associated with prior optimistic bias in financial reporting.

Our empirical analysis shows that the likelihood of reporting larger positive or smaller negative quarterly earnings surprises decreases with the beginning balance of net operating assets relative to sales, suggesting that managers' ability to optimistically bias earnings decreases with the extent to which net asset values are already overstated on the balance sheet. Sensitivity tests suggest that our findings are robust. Nevertheless, our evidence should be interpreted with caution given the limitations of our study. In particular, our study focuses on only one incentive to manage earnings—meeting or beating analysts' earnings forecasts. However, managers have other incentives to manage earnings, some of which may lead to pessimistically rather than optimistically biased earnings. Our study does not control for these other incentives. We leave this for future research.

APPENDIX

We derive Equation (2) by assuming that the cumulative probability of reporting an EPS surprise of less than k¢ is:

$$\Pr(\text{SURPRISE} < k' \mid \mathbf{x}) = F(-\mathbf{x}\boldsymbol{\beta}_k),$$

where $\mathbf{x}$ is a vector of independent variables, $\boldsymbol{\beta}_k$ is a vector of parameters for a predetermined earnings surprise benchmark k, and F is the cumulative logistic distribution:

$$F(-\mathbf{x}\boldsymbol{\beta}_k) = \exp(-\mathbf{x}\boldsymbol{\beta}_k)\ /\ [1 + \exp(-\mathbf{x}\boldsymbol{\beta}_k)].$$

To ensure that the sum of cumulative probabilities across all k equals 1, we impose the constraint $-\mathbf{x}\boldsymbol{\beta}_k \geq -\mathbf{x}\boldsymbol{\beta}_{k-1}$ for all k. The odds of reporting an earnings surprise of at least k¢ instead of less than k¢ are:

$$\begin{aligned} \Omega_k(\mathbf{x}) &= \Pr(\text{SURPRISE} \geq k \mid \mathbf{x})\ /\ \Pr(\text{SURPRISE} < k \mid \mathbf{x}) \\ &= [1 - F(-\mathbf{x}\boldsymbol{\beta}_k)]\ /\ F(-\mathbf{x}\boldsymbol{\beta}_k) = \exp(\mathbf{x}\boldsymbol{\beta}_k), \end{aligned}$$

which is a general version of Equation (2).

To determine the effect of a change in $\mathbf{x}$ on the odds of reporting an earnings surprise of at least k¢, suppose that $\mathbf{x}$ changes from $\mathbf{x} = \mathbf{x}_1$ to $\mathbf{x} = \mathbf{x}_2$. The odds then change from $\Omega_k(\mathbf{x}_1)$ to $\Omega_k(\mathbf{x}_2)$ by the factor:

$$\Omega_k(\mathbf{x}_2)\ /\ \Omega_k(\mathbf{x}_1) = \exp(\mathbf{x}_2\boldsymbol{\beta}_k)\ /\ \exp(\mathbf{x}_1\boldsymbol{\beta}_k) = \exp([\mathbf{x}_2 - \mathbf{x}_1]\boldsymbol{\beta}_k)$$

That is, the odds change by $100[\exp([\mathbf{x}_2 - \mathbf{x}_1]\boldsymbol{\beta}_k) - 1]$ percent. If only one variable, say x_j with parameter β_j, changes by δ, then the odds of reporting an earnings surprise of at least k¢ will change by $100[\exp(\delta \times \beta_{j,k}) - 1]$ percent, all else equal.

REFERENCES

Abarbanell, J. S., and V. L. Bernard. 1992. Tests of analysts' overreaction/underreaction to earnings information as an explanation for anomalous stock price behavior. *The Journal of Finance* 47 (July): 1181–1207.

———, and R. Lehavy. 2001a. Biased forecasts or biased earnings? The role of earnings management in explaining apparent optimism and inefficiency in analysts' earnings forecasts. Working paper, The University of North Carolina at Chapel Hill, NC.

———, and ———. 2001b. Differences in commercial database reported earnings: Implications for inferences concerning analyst forecast rationality, the association between prices and earnings, and firm reporting discretion. Working paper, The University of North Carolina at Chapel Hill, NC.

Accounting Principles Board (APB). 1973. *Interim Financial Reporting*. APB Opinion No. 28. New York, NY: APB.

Ali, A., and S. Kallapur. 2001. Securities price consequences to the Private Securities Litigation Reform Act of 1995 and related events. *The Accounting Review* 76 (July): 431–460.

Barth, M. E., J. A. Elliott, and M. W. Finn. 1999. Market rewards associated with patterns of increasing earnings. *Journal of Accounting Research* 37 (Autumn): 387–413.

Bartov, E., D. Givoly, and C. Hayn. 2002. The rewards to meeting or beating earnings expectations. *Journal of Accounting and Economics* 33 (June): 173–204.

Beatty, A., S. L. Chamberlain, and J. Magliolo. 1995. Managing financial reports of commercial banks: The influence of taxes, regulatory capital, and earnings. *Journal of Accounting Research* 33 (Autumn): 231–261.

Becker, C. L., M. L. DeFond, J. Jiambalvo, and K. R. Subramanyam. 1998. The effect of audit quality on earnings management. *Contemporary Accounting Research* 15 (Spring): 1–24.

Beneish, M. D. 1998. Discussion of: Are accruals during initial public offerings opportunistic? *Review of Accounting Studies* 3: 209–221.

Brant, R. 1990. Assessing proportionality in the proportional odds model for ordinal logistic regression. *Biometrics* 46 (December): 1171–1178.

Brown, L. D., and K-J. Kim. 1991. Timely aggregate analyst forecasts as better proxies for market earnings expectations. *Journal of Accounting Research* 29 (Autumn): 382–385.

———. 2001. A temporal analysis of earnings surprises: Profits versus losses. *Journal of Accounting Research* 39 (September): 221–241.

Burgstahler, D. C., and M. Eames. 2001. Management of earnings and analyst forecasts to achieve zero earnings surprise. Working paper, University of Washington, Seattle, WA.

Clogg, C. C., and E. S. Shihadeh. 1994. *Statistical Models for Ordinal Variables*. Thousand Oaks, CA: Sage.

Das, S., C. B. Levine, and K. Sivaramakrishnan. 1998. Earnings predictability and bias in analysts' earnings forecasts. *The Accounting Review* 73 (April): 277–294.

Dechow, P. M., R. G. Sloan, and A. P. Sweeney. 1995. Detecting earnings management. *The Accounting Review* 70 (April): 193–225.

———, ———, and ———. 1996. Causes and consequences of earnings manipulations: An analysis of firms subject to enforcement actions by the SEC. *Contemporary Accounting Research* 13 (Spring): 1–35.

———, and D. J. Skinner. 2000. Earnings management: Reconciling the views of accounting academics, practitioners, and regulators. *Accounting Horizons* 14 (June): 235–250.

Degeorge, F., J. Patel, and R. Zeckhauser. 1999. Earnings management to exceed thresholds. *The Journal of Business* 72 (January): 1–33.

Easterwood, J., and S. Nutt. 1999. Inefficiency in analysts' earnings forecasts: Systematic misreaction or systematic optimism? *The Journal of Finance* 54 (October): 1777–1797.

Elliott, J. A., and W. H. Shaw. 1988. Write-offs as accounting procedures to manage perceptions. *Journal of Accounting Research* 26 (Supplement): 91–119.

Ettredge, M. L., D. T. Simon, D. B. Smith, and M. S. Stone. 2000. The effect of the external accountant's review on the timing of adjustments to quarterly earnings. *Journal of Accounting Research* 38 (Spring): 195–207.

Feroz, E. H., K. Park, and V. S. Pastena. 1991. The financial and market effects of the SEC's accounting and auditing enforcement releases. *Journal of Accounting Research* 29 (Supplement): 107–142.

Fields, T. D., T. Z. Lys, and L. Vincent. 2001. Empirical research on accounting choice. *Journal of Accounting and Economics* 31 (September): 255–307.

Financial Accounting Standards Board (FASB). 1974. *Reporting Accounting Changes in Interim Financial Statements*. Statement of Financial Accounting Standards No. 3. Stamford, CT: FASB.

———. 1978. *Objectives of Financial Reporting by Business Enterprises*. Statement of Financial Accounting Concepts No. 1. Stamford, CT: FASB.

———. 1985. *Elements of Financial Statements*. Statement of Financial Accounting Concepts No. 6. Stamford, CT: FASB.

Francis, J., D. Philbrick, and K. Schipper. 1994. Shareholder litigation and corporate disclosures. *Journal of Accounting Research* 32 (Autumn): 137–164.

Fu, V. K. 1998. sg88: Estimating generalized ordered logit models. *Stata Technical Bulletin* 44: 27–30.

Gaver, J. J., and J. S. Paterson. 1999. Managing insurance company financial statements to meet regulatory and tax reporting goals. *Contemporary Accounting Research* 16 (Summer): 207–241.

Han, J. C. Y., and S-W. Wang. 1998. Political costs and earnings management of oil companies during the 1990 Persian Gulf crisis. *The Accounting Review* 73 (January): 103–117.

Hunt, A., S. E. Moyer, and T. Shevlin. 1996. Managing interacting accounting measures to meet multiple objectives: A study of LIFO firms. *Journal of Accounting and Economics* 21 (June): 339–374.

Johnson, N. S. 1999. Current SEC developments—"Managed earnings" and "The year of the accountant." Remarks delivered at the Utah State Bar Mid-Year Convention, St. George, UT. March 6. Available at: http://www.sec.gov/news/speech/speecharchive/1999/spch264.htm.

Jones, J. J. 1991. Earnings management during import relief investigations. *Journal of Accounting Research* 29 (Autumn): 193–228.

Levitt, A. 1998. The numbers game. Remarks delivered at New York University Center for Law and Business, New York, NY. September 28. Available at: http://www.sec.gov/news/speech/speecharchive/1998/spch220.txt.

Libby, R., and W. R. Kinney, Jr. 2000. Does mandated audit communication reduce opportunistic corrections to manage earnings to forecasts? *The Accounting Review* 75 (October): 383–404.

Lim, T. 2001. Rationality and analysts' bias. *The Journal of Finance* 56 (February): 369–385.

Lys, T., and R. L. Watts. 1994. Lawsuits against auditors. *Journal of Accounting Research* 32 (Supplement): 65–93.

Matsumoto, D. A. 2002. Management's incentives to avoid negative earnings surprises. *The Accounting Review* 77 (July): 483–514.

Mendenhall, R. R., and W. D. Nichols. 1988. Bad news and differential market reactions to announcements of earlier-quarters versus fourth-quarter earnings. *Journal of Accounting Research* 26 (Supplement): 63–86.

Moehrle, S. R. 2002. Do firms use restructuring charge reversals to meet earnings targets? *The Accounting Review* 77 (April): 397–413.

Nissim, D., and S. H. Penman. 2001. Ratio analysis and equity valuation: From research to practice. *Review of Accounting Studies* 6 (March): 109–154.

Palmrose, Z-V., V. J. Richardson, and S. Scholz. 2001. Determinants of market reactions to restatement announcements. Working paper, University of Southern California, Los Angeles, CA.

Payne, J. L., and S. G. Robb. 2000. Earnings management: The effect of ex ante earnings expectations. *Journal of Accounting, Auditing and Finance* 15 (Fall): 371–392.

Skinner, D. J., and R. G. Sloan. 2001. Earnings surprises, growth expectations, and stock returns or don't let an earnings torpedo sink your portfolio. Working paper, University of Michigan, Ann Arbor, MI.

Soffer, L. C., S. R. Thiagarajan, and B. R. Walther. 2000. Earnings preannouncement strategies. *Review of Accounting Studies* 5 (March): 5–26.

St. Pierre, K., and J. A. Anderson. 1984. An analysis of the factors associated with lawsuits against public accountants. *The Accounting Review* 59 (April): 242–263.

Teoh, S. H., T. J. Wong, and G. R. Rao. 1998. Are accruals during initial public offerings opportunistic? *Review of Accounting Studies* 3: 175–208.

THE ACCOUNTING REVIEW
Vol. 77
Supplement 2002
pp. 29–33

DISCUSSION OF
The Balance Sheet as an Earnings Management Constraint

Mark L. DeFond
University of Southern California

I. INTRODUCTION

Barton and Simko (2002) (hereafter the paper) contribute to the earnings management literature by suggesting that earnings management is constrained by the fundamental characteristic of double entry accounting that links the balance sheet with the income statement: managerial judgments and estimates that increase earnings will also increase net assets, but such increases are limited by the necessity to recognize and measure net assets in conformity with GAAP. The paper's evidence is essentially consistent with the view that GAAP-imposed constraints on asset and liability measurements put upper limits on the extent to which managers can opportunistically increase reported earnings. Relative to prior earnings management research, which tends to focus on the *income statement* implications of managed earnings and the role of forces such as auditing and litigation *outside* of the accounting process to constrain earnings management, this paper is distinguished by its focus on the balance sheet and the role of accounting standards as constraints on earnings management.

The paper also contributes to the earnings management literature by offering an explanation for why some firms fall short of analysts' forecasted earnings by as little as one cent per share. Explaining this behavior is important because several recent papers find that investors impose large penalties on firms that just miss analysts' forecasts (Skinner and Sloan 2001; Dechow et al. 2000; Bartov et al. 2002; DeFond and Park 2002). While intuition suggests that managers should be able to exercise their reporting judgment to boost earnings by one cent per share, the paper identifies circumstances where this is not the case.

An additional contribution of the paper is that the study identifies a setting where market participants are likely to find balance sheet information useful in interpreting quarterly earnings announcements. The paper's findings suggest that information in balance sheets may help investors understand why managers miss forecasted earnings. Identifying a demand for financial information to assist investors in interpreting quarterly earnings at the earnings announcement date is consistent with recent research that finds increased disclosure of a variety of financial information concurrently with earnings announcements (Francis et al. 2002; Chen et al. 2002). Although this is not the authors' intent, the paper's findings imply that managers may provide these additional disclosures in response to investor demand for information to supplement quarterly earnings announcements.

Finally, the paper's approach to investigating earnings management follows suggestions in recent articles that review and critique prior earnings management research. Specifically, the paper's primary analysis examines evidence of earnings management conditioned on *managerial incentives* to achieve *capital market-based* earnings targets, as opposed to accounting-based earnings targets. Dechow and Skinner (2000) and Healy and Wahlen (1999), among others, argue that tests of earnings management should be based on capital market incentives because managers are likely to be more strongly motivated by market incentives than by accounting-based incentives. This motivation is likely to be particularly strong during the period the paper analyzes (1993–1999) because of the emphasis on stock-based rewards to managers over this period (Dechow and Skinner 2000).

Overall, the authors address an interesting question, perform a careful analysis, and find some intuitive relations that add to the earnings management literature. Like all research, however, the paper has limitations. Sections II through IV of this discussion focus on problems of empirically capturing the concept represented by balance sheet constraints on earnings management and evaluate some of the paper's assumptions and inferences. Section V proposes some questions for future research that are suggested by the paper's findings.

II. EMPIRICAL MEASURES OF BALANCE-SHEET-BASED EARNINGS MANAGEMENT CONSTRAINTS

The Definition vs. the Empirical Measures

One limitation of the paper is that the empirical constructs used to capture the concept of earnings management constraints imposed by balance sheet measures of net assets do not precisely map into the paper's conceptual definition. The authors essentially argue that firms will encounter balance-sheet-based constraints on earnings management after a history of income-increasing earnings management because the past earnings management means their net assets are currently overstated. The intuition underlying this definition is straightforward, suggesting that managers who opportunistically boost earnings will eventually face new costs in the form of net assets whose balance sheet measures overstate future economic benefits to be realized. While the definition is reasonable and intuitive, the paper's empirical measures only loosely conform to this definition.

The paper captures the concept of net asset overstatements as net operating assets relative to sales (NOA) in the primary analysis in the paper's Table 5, and performs additional analysis in Table 6 that separately examines three components of NOA: net working capital scaled by sales (WC), net fixed assets scaled by sales (FIXED_ASSETS), and other long-term assets scaled by sales (OTHER_LTASSETS, which is defined as NOA minus WC and FIXED_ASSETS). However, none of these measures *necessarily* identifies either asset overstatements or the effects of past income-increasing earnings management. Rather, the measures capture whether net assets (relative to sales) tend to be relatively large compared to the population. While firms with larger ratios, other things equal, are probably more likely to have overstated net assets, and while Table 3 attempts to document that firms with larger ratios have histories of earnings management, the measures could be improved to capture more directly the construct implied by balance sheet constraints on earnings management.

A Model of "Expected" Net Assets to Sales Ratios as a Function of Industry and Firm-Specific Factors

The assets-to-sales ratios used by the authors to capture constraints on earnings management measure the magnitudes of various categories of assets-to-sales, but they provide no information on the *expected* magnitudes. This is an important limitation because the paper's arguments rely on firms facing constraints on earnings management when net assets to sales ratios are *unusually* large. The underlying (and not unreasonable) assumption on which the current measures are based is that relatively large ratios imply a greater likelihood of overstated assets. At a minimum, however, there are likely to be systematic differences in the ratios across industries, as well as firm-specific effects on the ratios, that are unrelated to whether net assets are overstated. A properly specified expectations model would reduce the noise in measuring overstated assets, yielding a more precise empirical construct that would facilitate a more powerful research design.

Recent research provides some direction for developing an expectations model for ratios of net assets to sales. Dechow et al. (1998) model the relation of working capital to sales by assuming that normal working capital is a fixed proportion of contemporaneous sales, and that sales follow a random walk. This assumption is consistent with empirical observations that find that the ratio of net working capital to sales among Compustat firms is relatively constant over time, suggesting that expected or normal net working capital is a function of current period sales. DeFond and Park (2001) use this observation to develop a simple model of expected working capital as a function of seasonally adjusted quarterly sales. In the context of the earnings management question addressed in the paper, both firm-specific and industry factors could be incorporated in developing an expectations model for net asset measures.

The Relation between Firm Performance and Measures of Balance Sheet Constraints on Earnings Management

One Conference participant noted that the net-assets-to-sales measures used in the paper are inverse asset turnover ratios. Turnover ratios, which provide information on how efficiently managers employ the firm's assets, are intended to capture some aspects of firm performance. Since the ability to meet analysts' forecasts is also related to performance, the associations between the net asset ratios and the ability to meet forecasts might actually be driven by differences in firm performance, partly or entirely unrelated to constraints imposed by the accumulated effects, in balance sheet measures of past earnings management. The authors appropriately address this issue by including several control variables that capture performance in their multivariate tests.

While the paper's multivariate tests control for firm performance, the univariate analysis designed to test the assumption that firms with relatively large ratios of net operating assets to sales have a history of earnings management does not control for performance. Table 3 in the paper reports that firms in the *top* quintile of net operating assets to sales (within their industry) tend to report Jones (1991) model-type working capital discretionary accruals that are relatively more income-increasing over the prior 20 quarters, when compared to firms in the *bottom* quintile of net operating assets to sales. However, Jones model-type discretionary accruals are known to be mechanically associated with performance (Dechow et al. 1995). Because both Jones model discretionary accruals and asset turnover ratios capture aspects of firm performance, the associations shown in Table 3 may be unrelated to past earnings management effects on balance sheet measures. Rather, Table 3 may document a relation induced by firm performance.

Another limitation of the analysis in Table 3 is that the discretionary accruals in that analysis are not conditioned on managerial incentives to increase earnings. Without such conditioning, it is difficult to infer earnings management (Dechow and Skinner 2000). For example, because Jones model-type discretionary accruals are correlated with performance, the pattern of past income-increasing discretionary accruals in Table 3 might be interpreted as a demonstration that firms with high ratios of net assets to sales have a history of superior performance (because income-increasing discretionary accruals are correlated with high reported earnings). To convincingly infer a history of income-increasing earnings management from Table 3, the authors need to demonstrate that the discretionary accruals pattern is consistent with management meeting earnings targets such as analysts' forecasts. This demonstration would also make the analysis in Table 3 consistent with the approach used in the paper's multivariate tests.

III. DOES THE ANALYSIS NEED TO ASSUME A HISTORY OF PAST EARNINGS MANAGEMENT?

The prior discussion argues that confounding factors make it difficult to conclude from Table 3 that firms with relatively large net-operating-assets-to-sales ratios necessarily have a history of earnings management. A more fundamental issue is whether this conclusion is really necessary for the hypothesis addressed in this paper. One Conference participant pointed out that the authors' conclusions regarding the role of the balance sheet in constraining earnings management do not require a history of income-increasing earnings management. That is, *how* the balance sheet becomes a constraint on earnings management is irrelevant to the conclusion that firms with large ratios of assets to sales in the current period find it harder to manage earnings in the current period. The asset overstatements could result from past earnings management (which implies management deception) or from honest errors in managerial judgments and estimates. Dropping the requirement that large ratios of assets to sales must arise from past earnings management puts less of a burden on the paper's empirical analysis without changing the paper's overall contribution.

IV. CAN THE BALANCE SHEET ALSO *FACILITATE* EARNINGS MANAGEMENT?

The authors emphasize that constraints imposed by the accumulated effects of past earnings management in balance sheet accounts reduce the ability to manage earnings upward; one implication is that unusually *low* net asset values effectively *facilitate* income-increasing earnings management. This implication is consistent with the analysis in Table 5. In a regression of the authors' measure of earnings surprise on several determinants of earnings surprise, Panel A of Table 5 reports that the coefficients on the net-operating-assets-to-sales ratio are significantly negative. This result indicates that the odds of beating market expectations are *greater* for firms with *smaller* net-operating-assets-to-sales ratios. This suggests that firms with relatively low net operating asset values find it *easier* to boost earnings through income-increasing discretionary accounting choices. Thus, while the paper portrays the reversing feature of accruals as *deterring* earnings management, this same characteristic of the accounting process is also likely to *facilitate* income-increasing earnings management among firms with small net operating assets to sales ratios.

The paper also notes that relatively larger net-assets-to-sales ratios deter only income-*increasing* earnings management. Consistently, the regression results in Table 5 indicate that balance sheet constraints on income-increasing earnings management actually facilitate income-*decreasing* earnings management. That is, Panel B of Table 5 reports that firms with relatively larger values of net operating assets to sales are less likely to exceed earnings expectations by a large amount, consistent with these firms having greater discretion in reducing current earnings. The converse is implied for firms with relatively smaller net asset ratios. That is, relatively smaller net assets to sales appear to deter income-*decreasing* earnings management. That said, the paper's focus on constraints on income-increasing earnings management seems appropriate since, arguably, regulators and auditors appear to find income-increasing earnings management most objectionable.

V. FUTURE RESEARCH

The paper raises several issues that could be explored in future research, some of which were hinted at in the Conference discussion. For example, managers may use asset impairment charges to avoid balance sheet constraints on earnings management and thus improve their ability to meet or beat analysts' forecasts. This raises the question of why any managers would allow their firm's net asset values to become relatively large. Thus, one question posed by the paper's findings turns on the trade-offs between using (income-decreasing) asset impairment charges to preserve earnings management flexibility and suffering the penalties associated with missing analysts' forecasts.

In addition, future research might explore the paper's implication that market participants are likely to find balance sheet information useful in assessing the existence and/or extent of earnings management. This implies that market participants, at least in some settings, are likely to demand balance sheet disclosures concurrently with earnings announcements. While Francis et al. (2002) and Chen et al. (2002) provide some initial investigation of disclosures that accompany earnings announcements, the findings in this paper suggest additional questions. For example, future researchers might explore whether disclosures that accompany earnings announcements are related to market participants' concerns about various aspects of managed earnings, and whether such disclosures, in turn, are related to the market's response to earnings announcements.

REFERENCES

Barton, J. J., and P. J. Simko. 2002. The balance sheet as an earnings management constraint. *The Accounting Review* (Supplement): 1–27.

Bartov, E., D. Givoly, and C. Hayn. 2002. The rewards to meeting or beating earnings expectations. *Journal of Accounting and Economics* 33 (June): 173–204.

Chen, S., M. DeFond, and C. W. Park. 2002. Voluntary disclosure of balance sheet information in quarterly earnings announcements. *Journal of Accounting and Economics* 33 (June): 229–251.

Dechow, P., R. Sloan, and A. Sweeney. 1995. Detecting earnings management. *The Accounting Review* 70 (April): 193–225.

———, S. P. Kothari, and R. Watts. 1998. The relation between earnings and cash flows. *Journal of Accounting and Economics* 25 (May): 133–168.

———, S. Richardson, and A. I. Tuna. 2000. Are benchmark beaters doing anything wrong? Working paper, University of Michigan.

———, and D. Skinner. 2000. Earnings management: Reconciling the views of academics, practitioners and regulators. *Accounting Horizons* 14 (June): 235–250.

DeFond, M., and C. W. Park. 2001. The reversal of abnormal accruals and the market valuation of earnings surprises. *The Accounting Review* 76 (July): 375–404.

———, and ———. 2002. The use of accounting discretion in managing earnings to meet market expectations. Working paper, University of Southern California and Hong Kong University of Science and Technology.

Francis, J., K. Schipper, and L. Vincent. 2002. Earnings announcements and competing information. *Journal of Accounting and Economics* 33 (August): 313–342.

Healy, P., and J. Wahlen. 1999. A review of the earnings management literature and its implications for standard setting. *Accounting Horizons* 13 (December): 365–383.

Jones, J. 1991. Earnings management during import relief investigations. *Journal of Accounting Research* 29 (Autumn): 193–228.

Skinner, D., and R. Sloan. 2001. Earnings surprises, growth expectations and stock returns, or don't let an earnings torpedo sink your portfolio. Working paper, University of Michigan, Ann Arbor, MI.

THE ACCOUNTING REVIEW
Vol. 77
Supplement 2002
pp. 35–59

The Quality of Accruals and Earnings: The Role of Accrual Estimation Errors

Patricia M. Dechow
University of Michigan
Ilia D. Dichev
University of Michigan

ABSTRACT: This paper suggests a new measure of one aspect of the quality of working capital accruals and earnings. One role of accruals is to shift or adjust the recognition of cash flows over time so that the adjusted numbers (earnings) better measure firm performance. However, accruals require assumptions and estimates of future cash flows. We argue that the quality of accruals and earnings is decreasing in the magnitude of estimation error in accruals. We derive an empirical measure of accrual quality as the residuals from firm-specific regressions of changes in working capital on past, present, and future operating cash flows. We document that observable firm characteristics can be used as instruments for accrual quality (e.g., volatility of accruals and volatility of earnings). Finally, we show that our measure of accrual quality is positively related to earnings persistence.

Keywords: *quality of accruals; earnings quality; estimation errors; earnings persistence.*

Data Availability: *Data are available from sources identified in the paper.*

I. INTRODUCTION

This paper suggests a new measure of one aspect of the quality of working capital accruals and earnings. The measure is based on the observation that accruals shift or adjust the recognition of cash flows over time, so that the adjusted numbers (earnings), better measure firm performance (e.g., see Statement of Accounting Concepts

---

We thank workshop participants at The University of Arizona, University of Chicago, Columbia University, Indiana University, University of Michigan, *The Accounting Review* Conference on Quality of Earnings held at Emory University, the 2001 Big Ten Research Conference, and especially Dan Collins, Bob Lipe, Mort Pincus, Katherine Schipper, Peter Wysocki, and two anonymous reviewers. Professor Dechow acknowledges financial support provided by the Michael A. Sakkinen Research Scholar Fund at the University of Michigan Business School. Professor Dichev acknowledges the financial support of the Sanford Robertson Assistant Professorship in Business Administration and the Norman Auerbach PricewaterhouseCoopers Faculty Fellowship.

No. 1, FASB 1978, para. 44). For example, recording a receivable accelerates the recognition of a future cash flow in earnings, and matches the timing of the accounting recognition with the timing of the economic benefits from the sale. However, accruals are frequently based on assumptions and estimates that, if wrong, must be corrected in future accruals and earnings. For example, if the net proceeds from a receivable are less than the original estimate, then the subsequent entry records both the cash collected and the correction of the estimation error. We argue that estimation errors and their subsequent corrections are noise that reduces the beneficial role of accruals. Therefore, the quality of accruals and earnings is decreasing in the magnitude of accrual estimation errors. Our empirical measure of accrual quality is the extent to which working capital accruals map into operating cash flow realizations, where a poor match signifies low accrual quality.

The intuition behind this measure is available in both theoretical and practical texts that observe that the beneficial role of accruals is reduced by various limitations, including estimation errors. For example, Palepu et al. (2000) discuss estimation errors as a factor that reduces accounting quality, and suggest that estimation accuracy depends on firm characteristics like complexity of transactions and predictability of the firm's environment. We develop this existing intuition into a practical measure of accrual and earnings quality.

Our approach to estimating the quality of accruals is related to two streams of research. First, several studies document the benefits of the accrual process, finding that earnings is a better measure of performance than the underlying cash flows (e.g., Dechow 1994; Dechow et al. 1998; Liu et al. 2002). We build on this evidence by exploring the tradeoffs inherent in the accrual process. Second, a number of studies use models of "discretionary accruals" to investigate the manipulation of accruals to achieve earnings management goals (see Healy and Wahlen [1999] for a recent review). These studies focus on the opportunistic use of accruals to window-dress and mislead users of financial statements. This stream of research suggests that managerial intent affects the incidence and magnitude of accrual estimation errors.

In contrast, we argue that even in the absence of intentional earnings management, accrual quality will be systematically related to firm and industry characteristics. This distinction is important because such characteristics are likely to be both observable and recurring (e.g., the volatility of operations is systematically related to the propensity to make estimation errors) as compared to the determinants of managerial opportunism that are often unobservable and/or sporadic (e.g., before stock offerings). For our purposes, we do not attempt to disentangle "intentional" estimation errors from unintentional errors because both imply low-quality accruals and earnings.

We develop a model that examines the origination and reversal of working capital accruals in a stylized firm. The model embodies the intuition that the timing of the firm's economic achievements and sacrifices often differs from the timing of the related cash flows, and that the benefit of accruals is to adjust for these cash flow timing problems. However, the model also reveals that the benefit of using accruals comes at the cost of including accrual components that initiate and correct estimation errors. We focus on working capital accruals and operating cash flows for tractability: the initiation and reversal of these accruals occurs within a year. Our measure of accrual estimation errors is the residuals from firm-specific regressions of changes in working capital on last year, present, and one-year ahead cash flows from operations. These residuals are unrelated to cash flow realizations, and include the estimation errors and their reversals. The standard deviation of these residuals is our firm-specific measure of quality of accruals and earnings, where a higher standard deviation signifies lower quality.

We illustrate the usefulness of our analysis in two ways. First, we explore the relation between our measure of accrual quality and firm characteristics. The nature of the accrual process suggests that the magnitude of estimation errors will be systematically related to business fundamentals like the length of the operating cycle and variability of operations. We find that accrual quality is negatively related to the absolute magnitude of accruals, the length of the operating cycle, loss incidence, and the standard deviation of sales, cash flows, accruals, and earnings, and positively related to firm size. Our results suggest that these observable firm characteristics can be used as instruments for accrual quality. This is important because the regression-based estimation of accrual quality demands long time-series of data and the availability of subsequent cash flows, which makes it costly or infeasible for certain practical applications (e.g., quality-of-accruals-based trading strategies).

Second, we illustrate the usefulness of our analysis by exploring the relation between our measure of accrual quality and earnings persistence. Firms with low accrual quality have more accruals that are unrelated to cash flow realizations, and so have more noise and less persistence in their earnings. Indeed, we find a strong positive relation between accrual quality and earnings persistence. However, our measure of accrual quality is theoretically and empirically related to the absolute magnitude of accruals, and Sloan (1996) documents that the level of accruals is less persistent than cash flows. Probing further, we find that accrual quality and level of accruals are incremental to each other in explaining earnings persistence, with accrual quality the more powerful determinant.

The remainder of the study is organized as follows. Section II presents our model of accrual quality. Section III describes the sample and provides descriptive statistics. Section IV derives the empirical measure of accrual quality and explores the relation between accrual quality and both firm characteristics and earnings persistence. Our conclusions are provided in Section V.

II. THEORETICAL MEASURE OF ACCRUAL AND EARNINGS QUALITY

Model of Accruals

Our model of accruals focuses on working capital accruals because cash flow realizations related to working capital generally occur within one year, making both the theory and the empirics more tractable. While the intuition about errors in estimation applies to all accruals, the long lags between noncurrent accruals and cash flow realizations practically restrict the application of our approach to only short-term accruals.

We build our accrual framework around the observation that earnings equals cash flows plus accruals, E = CF + Accruals. Cash flows for any period t can be categorized into three groups:

Symbol	Name
CF_t^{t-1}	Cash Collections or Payments of Amounts Accrued at t−1 (net)
CF_t^t	Current Cash Flows (net)
CF_t^{t+1}	Cash Flows Deferred to t+1 (net)

From an accounting perspective, for each cash flow, the two important events are the receipt or disbursement of the cash flow and the recognition of that cash flow in earnings (as revenue or expense). The subscript refers to the period the cash flow is received or disbursed, and the superscript refers to the period the cash flow is recognized in earnings. For example, CF_t^{t-1} denotes that the cash flow occurs after the corresponding amount is recognized in earnings (e.g., collection of an accounts receivable). CF_t^t refers to cash flows received or paid in the same period as the cash flows are recognized in earnings. Finally, CF_t^{t+1} refers to cash received or paid before the revenue or expense is recognized in earnings, such as cash payments for inventory.

Summing up, total cash flow for period t is:

$$CF_t = CF_t^{t-1} + CF_t^{t} + CF_t^{t+1} \tag{1}$$

The accounting system provides for accruals, temporary adjustments that shift the recognition of cash flows over time. When recognition of a cash flow is shifted, two accrual entries are created, an opening and a closing accrual. The opening accrual is initiated when either (1) revenue or expense is recognized before the cash is received or paid or (2) cash is received or paid before it is recognized in earnings. The closing accrual is recorded when the other element of this pair has occurred and reverses the accrual portion of the original entry.

When cash flows occur *after* the corresponding revenues and expenses are recognized in earnings, managers must estimate the amount of cash to be received or paid in the future. To the extent that cash flow realizations differ from their accrual estimates, the opening accrual will contain an estimation error that is corrected by the closing accrual.

We incorporate this intuition in our model using the following notation for opening and closing accruals related to future cash flows:

Name	**Accrual**	**= Amount**
Accrual for future collections and payments—Opening Same sign as related cash flow	$A_{CFt+1/t}^{O}$	$= CF_{t+1}^{t} + \varepsilon_{t+1}^{t}$
Accrual for future collections and payments—Closing Opposite sign of related cash flow	$A_{CFt/t-1}^{C}$	$= -CF_t^{t-1} - \varepsilon_t^{t-1}$

For accruals, the subscript refers to the corresponding cash flow, while the superscript indicates opening or closing accruals. For example, $A_{CFt+1/t}^{O}$ is an opening accrual related to cash flow CF_{t+1}^{t}. The opening accrual at time t, which reflects the expectation of the t+1 cash flow, is equal to the actual t+1 cash flow plus an error term that reflects the difference between the accrual expectation and the cash flow realization. We assume that all accruals are resolved within one period so the closing accrual for cash collections and payments at time t offsets the opening accrual from t−1. The closing accrual is equal to the actual cash flow collected or paid in t plus an error term equal to the difference between last period's expectation and this period's cash flow realization. Thus, each period's total accruals contain an estimation error in the opening accruals (the actual value of which is determined next period, ε_{t+1}^{t}) plus a realized error in the closing accruals (determined from current period cash flow realizations, ε_t^{t-1}).[1]

We next provide the notation for situations where cash flows are received or paid *before* their recognition in earnings. In this case, the accounting system recognizes the amount of the cash flow either as deferred revenue (for cash inflows) or a deferred cost (for cash outflows):

---

[1] We assume that estimation errors are independent of each other and of the magnitude of the cash flow realizations, *within the firm*. While estimation errors can be correlated across firms (e.g., bad debt expense proves to be systematically understated in a recession), cross-firm dependencies are not a problem for our approach because our tests are at the firm level. Our approach also ignores estimation errors that are correlated at the firm level (e.g., an optimistic manager systematically underestimates expenses). This is one of the limitations of our approach.

Name	Accrual	= Amount
Accrual that defers the recognition of cash flows—Opening Opposite sign of related cash flow	$A_{CFt/t+1}{}^{O}$	$= -CF_t{}^{t+1}$
Accrual that defers the recognition of cash flows—Closing Same sign as related cash flow	$A_{CFt-1/t}{}^{C}$	$= CF_{t-1}{}^{t}$

As before, accrual subscripts refer to the corresponding cash flows, while superscripts indicate opening or closing accruals. For example, $A_{CFt/t+1}{}^{O}$ is an opening accrual related to cash flow $CF_t{}^{t+1}$. Since deferred cash flows occur *before* recognition, these accruals contain no estimation error (in the sense of our model). For example, an inventory purchase initiates a positive accrual equal to the purchase price cash outflow, while the sale of the inventory triggers a negative accrual equal in magnitude to the initial cash flow. Appendix A provides a discussion and examples of whether and how our model detects accrual estimation errors depending on the timing of related accruals and cash flows.

Using Expression (1), and defining total accruals as the sum of opening and closing accruals, allows us to express earnings as:

$$E_t = CF_t + Accruals_t$$

$$E_t = (CF_t{}^{t-1} + CF_t{}^{t} + CF_t{}^{t+1}) + (A_{CFt+1/t}{}^{O} + A_{CFt/t-1}{}^{C} + A_{CFt/t+1}{}^{O} + A_{CFt-1/t}{}^{C})$$

$$E_t = (CF_t{}^{t-1} + CF_t{}^{t} + CF_t{}^{t+1}) + (CF_{t+1}{}^{t} + \varepsilon_{t+1}{}^{t} - CF_t{}^{t-1} - \varepsilon_t{}^{t-1} - CF_t{}^{t+1} + CF_{t-1}{}^{t}) \quad (2)$$

Re-arranging yields:

$$E_t = CF_{t-1}{}^{t} + CF_t{}^{t} + CF_{t+1}{}^{t} + \varepsilon_{t+1}{}^{t} - \varepsilon_t{}^{t-1}. \quad (3)$$

Equation (3), which presents earnings as the sum of past, present, and future cash flows plus an adjustment for estimation errors and their correction, expresses the main intuition of the paper. The intertemporal shifting of cash flows reflected in the cash flow terms of Equation (3) alleviates the timing problems of using current cash flows as a measure of performance. However, this benefit comes at the cost of using estimates, with the result that earnings includes both errors in estimates and their corrections. Both errors and error corrections reduce the quality of earnings as a measure of performance.

Most of our predictions about the relations between the variables in Equations (2) and (3) concern relations studied by existing research. We include them to show that our model embodies the intuitive and well-known properties of earnings, cash flows, and accruals, and as a benchmark and link to existing results. Using Equations (2) and (3) allows one to make the following predictions:

a) $Corr(E_t, CF_t) = +$
b) $Corr(E_t, Accruals_t) = +$
c) $Corr(CF_t, Accruals_t) = -$

These predictions, concerning the correlations between contemporaneous earnings, cash flows, and accruals, have been confirmed by, for example, Dechow (1994) and Dechow et al. (1998).

d) $\text{Corr}(E_t, CF_{t+1}) = +$
e) $\text{Corr}(\text{Accruals}_t, CF_{t+1}) = +$

These correlations suggest that earnings, and particularly its accrual portion, anticipates future cash flows. This result has been confirmed by, for example, Finger (1994) and Barth et al. (2001). Note that these predictions are based on *ceteris paribus* assumptions, so there is a need to control for possible confounding factors. Specifically, in examining the relation between current accruals and future cash flows, one needs to control for the confounding effect of current cash flows. The reason is that current accruals are negatively related to current cash flows, and future cash flows are (empirically) positively related to current cash flows.

f) $\text{Corr}(E_t, CF_{t-1}) = +$
g) $\text{Corr}(\text{Accruals}_t, CF_{t-1}) = +$

To our knowledge, the prediction that current earnings, and particularly its accrual portion, is positively correlated with *past* cash flows is novel. The intuition is that accruals defer the recognition of some past cash flows into current earnings. As mentioned above, in examining this prediction, it is important to control for the confounding effect of current cash flows.

Deriving an Empirical Measure of Quality of Accruals

We obtain an expression for accruals at time t by rearranging the accrual portion of earnings in Equation (2):

$$A_t = CF_{t-1}{}^{t} - (CF_t^{t+1} + CF_t^{t-1}) + CF_{t+1}{}^{t} + \varepsilon_{t+1}{}^{t} - \varepsilon_t^{t-1}. \tag{4}$$

Equation (4) conveys that (1) accruals are temporary adjustments that delay or anticipate the recognition of realized cash flows plus an estimation error term; (2) accruals are negatively related to current cash flows and positively related to past and future cash flows; and (3) the error term captures the extent to which accruals map into cash flow realizations, and can be used as a measure of accrual and earnings quality.

To derive practical measures of working capital accrual quality, we use the following firm-level time-series regression:

$$\Delta WC_t = b_0 + b_1{*}CFO_{t-1} + b_2{*}CFO_t + b_3{*}CFO_{t+1} + \varepsilon_t. \tag{5}$$

Relative to Equation (4), our measure of accruals is changes in working capital, and our proxies for cash flows related to accruals is cash flow from operations (CFO). The residuals from the regression reflect the accruals that are unrelated to cash flow realizations, and the standard deviation of these residuals is a firm-level measure of accrual quality, where higher standard deviation denotes lower quality.

The theoretical specification in Equation (4) uses *only* the portions of past, present, and future cash flows that are related to *current* accruals. Since we cannot identify these cash flow components, the empirical version in Expression (5) uses total CFOs.[2] Thus, the independent variables in Expression (5) are measured with error, implying that the regression coefficients are likely to be biased toward 0, and the R^2 will be reduced. Since the theoretical values of the coefficients from Equation (4) are $b_1 = 1$, $b_2 = -1$, and $b_3 = 1$, we expect that $0 < b_1 < 1$, and $-1 < b_2 < 0$, and $0 < b_3 < 1$. Appendix B provides a more expansive treatment of this issue, including simulation results that confirm that the estimated coefficients and the R^2 will be biased toward zero. Appendix B also reveals that the bias is more severe for b_1 and b_3 as compared to b_2, because the error component in past and future cash flows is greater than the component in current cash flows. Later in the paper (Section IV), we implement additional tests to control for the effect of this measurement error on our results. Finally, we add an intercept to Expression (5) to capture possible nonzero average accruals (e.g., average positive working capital accruals due to firm growth).

III. SAMPLE SELECTION, DESCRIPTIVE STATISTICS, AND CALIBRATION TESTS

Table 1 summarizes our sample selection. Our sample is obtained from the Compustat annual industrial and research files over 1987 to 1999. Given Collins and Hribar's (2002) result that the balance-sheet approach to deriving CFO leads to noisy and biased estimates, we use CFO as given in the Statement of Cash Flows reported under the Statement of Financial Accounting Standards No. 95 (SFAS No. 95, FASB 1987).[3] The sample is restricted to firms with complete data for assets, earnings, cash flow from operations, changes in accounts receivable, and changes in inventory. The last two requirements ensure sample firms have significant working capital accruals. Truncating the most extreme 1 percent of cash from operations, earnings, and changes in working capital, and requiring at least one

---

[2] Property-casualty insurance firms provide information about their accrual estimates, the subsequent cash flow realizations, and the resulting estimation errors (e.g., Petroni 1992; Anthony and Petroni 1997; Beaver and McNichols 1998). We do not pursue an industry-specific analysis because the *cross-industry* variation in the magnitude of estimation errors is likely to be larger than the typical within-industry variation, and because our interest is in more generalizable results.

[3] This Standard required the presentation of the statement of cash flows for fiscal years ending after July 15, 1988. However, some firms adopted this standard early in 1987.

TABLE 1
Derivation of Sample, 1987 to 1999

Firm-years with available cash from operations, earnings, and changes in accounts receivable and inventory and total assets reported in the statement of cash flows	59,360
Firm-years after truncation of the most extreme 1 percent of earnings, cash from operations, and changes in working capital	55,850
Firm-years with both lead and lag values of cash from operations	30,317
Firms with 8 or more annual observations	1,725
Firm-years available for the 1,725 firms used in the remaining analysis	15,234
Three-digit SIC groups with over 50 observations	136
Firm-years available for the 136 industries	27,204

All data is from Compustat.

year of past and future cash flows and earnings, yields a sample of 30,317 firm-years. In addition, we require at least eight years of data to estimate firm-specific regressions. This restriction reduces our sample to 15,234 firm-year observations for 1,725 firms.[4] For some of our industry-level analyses, less stringent data requirements yield a sample of 27,204 firm-year observations over 136 three-digit SIC industries.

Cash flow from operations is Compustat item 308. The change in working capital from year $t-1$ to t (ΔWC), is computed as ΔAR + ΔInventory − ΔAP − ΔTP + ΔOther Assets (net), where AR is accounts receivable, AP is accounts payable, and TP is taxes payable. Specifically, ΔWC is calculated from Compustat items as $\Delta WC = -$ (item 302 + item 303 + item 304 + item 305 + item 307).

We calculate earnings after short-term accruals but before long-term accruals (Earn) as Earn = CFO + ΔWC. We report Earnings before extraordinary items (Prof) as Compustat item 123, and Accruals (Prof-CFO) to provide comparability with other research. All variables are scaled by average total assets. We also calculate the length of the operating cycle (OC) as:

$$360/(\text{Sales}/\text{Average AR}) + 360/(\text{Cost of Goods Sold})/(\text{Average Inventory})$$

where Sales is Compustat item 12, cost of goods sold is item 41, AR is item 2, and Inventory is item 3.

Descriptive statistics and correlations are provided in Table 2. An examination of Panel A reveals that descriptive statistics are in line with those of other studies using similar variables and time period (e.g., Barth et al. 2001). Earn exceeds CFO, implying that short-term accruals are mostly positive. This is not surprising, given that most firms are growing and therefore increasing their working capital. As expected, average accruals are negative (−0.046), primarily because of depreciation.

The Pearson correlations in Panels B and C of Table 2 illustrate the relations between our sample variables and provide comparability with previous research (results for Spearman correlations are similar). These empirical correlations are in agreement with existing findings and the predictions of the model. Specifically, there is a positive contemporaneous correlation between Earn and CFO (0.73), and between Earn and ΔWC (0.33), and a negative correlation between CFO and ΔWC (−0.41). We also find that Accruals and ΔWC are highly positively correlated (0.75), suggesting that working capital accruals capture much of the variation in total accruals.

Consistent with Barth et al. (2001), we find that earnings and changes in working capital anticipate future cash flows from operations. Note that the simple correlation between ΔWC_t and CFO_{t+1} is −0.01 and statistically insignificant. As previously discussed, the reason is that ΔWC_t is negatively correlated with CFO_t, and CFO_t is positively correlated with CFO_{t+1} (0.56), which counteracts the expected positive relation between ΔWC_t and CFO_{t+1}. In Panel C of Table 2, we report that the partial correlation between ΔWC_t and CFO_{t+1}, controlling for CFO_t, is 0.29 (significant at the 0.0001 level). We also find that working capital accruals are positively related to past cash flows, implying that accruals defer the recognition of some past cash flows. The simple correlation between ΔWC_t and CFO_{t-1} in Panel B is only 0.008 but the partial correlation controlling for CFO_t in Panel C is 0.31 (significant at the 0.0001 level). Summarizing, our descriptive statistics and correlation results are in line with predictions and existing results, indicating that our model captures reasonably well some of the key features of accrual accounting.

---

[4] 481 firms have eight available observations, 1,054 have nine observations, and 190 firms have ten observations.

TABLE 2
Descriptive Statistics and Correlations for 15,234 Firm-Year Observations 1987 to 1999

Panel A: Descriptive Statistics

	Mean	*Standard Deviation*	*Lower Quartile*	*Median*	*Upper Quartile*
Cash flow from operations (CFO)	0.075	0.097	0.033	0.082	0.131
Change in working capital (ΔWC)	0.015	0.070	−0.015	0.010	0.044
Earnings before long-term accruals (Earn)	0.091	0.093	0.057	0.096	0.141
Earnings before extraordinary items (Prof)	0.030	0.113	0.009	0.042	0.081
Accruals (Prof-CFO)	−0.046	0.098	−0.084	−0.044	−.0035
Total Assets (in millions)	2,436	10,878	50	240	1,215

Panel B: Pearson Correlations

	$Earn_t$	CFO_t	ΔWC_t	CFO_{t+1}	CFO_{t-1}	$Earn_{t+1}$	$Accruals_t$
$Prof_t$	0.816*	0.575*	0.292*	0.458*	0.470*	0.540*	0.585*
$Earn_t$		0.728*	0.325*	0.571*	0.575*	0.674*	0.220*
CFO_t			−0.411*	0.558*	0.549*	0.597*	−0.327*
ΔWC_t				−0.011	0.008	0.072*	0.745*
CFO_{t+1}					0.482*	0.735*	−0.024
CFO_{t-1}						0.491*	−0.001
$Earn_{t+1}$							0.031*

Panel C: Partial Correlations (controlling for the effect of CFO_t)

Pearson Correlation	CFO_{t+1}	CFO_{t-1}
ΔWC_t	0.289*	0.306*

---

* Significant at the 0.0001 level.

Variable definitions:

Cash flow from operations (CFO) = item 308 from the Compustat Statement of Cash Flows;
Change in working capital (ΔWC) = ΔAR + ΔInventory − ΔAP − ΔTP + ΔOther Assets (net), where AR is accounts receivable, AP is accounts payable, and TP is taxes payable;
Earnings before long-term accruals (Earn) = CFO + ΔWC;
Earnings before extraordinary items (Prof) = Compustat item 123; and
Accruals = Prof − CFO.

All variables are scaled by average total assets.

IV. RESULTS

An Empirical Measure of Accrual Quality

Table 3 presents results of regressions of working capital accruals on past, present, and future cash flows from operations. First, we present firm-level regressions (Panel A) because

TABLE 3
Regressions of the Change in Working Capital on Past, Current, and Future Cash Flow from Operations for Firm-Years between 1987 to 1999

$$\Delta WC_t = b_0 + b_1 CFO_{t-1} + b_2 CFO_t + b_3 CFO_{t+1} + \varepsilon_t$$

	Intercept	b_1	b_2	b_3	*Adjusted* R^2
Panel A: Firm-Specific Regressions (1,725 firms)					
Mean	0.04	0.17	−0.62	0.09	0.47
(t-statistic)	(23.03)	(19.38)	(−57.06)	(10.38)	
Lower quartile	0.001	−0.02	−0.91	−0.10	0.23
Median	0.04	0.14	−0.65	0.09	0.55
Upper quartile	0.08	0.35	−0.35	0.28	0.80
Panel B: Industry-Specific Regressions (136 industries)					
Mean	0.03	0.19	−0.51	0.15	0.34
(t-statistic)	(16.09)	(21.10)	(−35.77)	(15.33)	
Lower quartile	0.01	0.11	−0.63	0.08	0.22
Median	0.03	0.18	−0.52	0.15	0.34
Upper quartile	0.04	0.26	−0.40	0.23	0.45
Panel C: Pooled Regression (15,234 firm-year observations)					
Coefficient	0.03	0.19	−0.51	0.18	0.29
(t-statistic)	(39.43)	(32.12)	(−78.29)	(29.18)	

The t-statistics in Panel A are determined based on the distribution of the 1,725 coefficients obtained from the firm-specific regressions requiring a minimum of eight observations per firm. T-statistics in Panel B are determined based on the distribution of the 136 coefficients obtained from three-digit SIC grouping regressions requiring a minimum of 50 observations per grouping.
All variables are defined as in Table 2.

our theory is defined and most naturally applied on a firm-level basis. In addition, we expect that a firm-level specification is superior to cross-sectional specifications because the regression coefficients are likely to differ across firms. The reason is that our operating cash flow proxies contain measurement error, which is likely to be systematically related to firm characteristics. For example, a long operating cycle implies that substantial future cash flows are recognized in current earnings (e.g., by accruing receivables). As a result, when future cash flows related to present accruals are a large part of total future cash flows, we expect that total future cash flows will load with a larger coefficient in Expression (5). Since such effects are likely related to firm-level characteristics, we believe that a firm-level specification is superior. However, we also present industry-specific and pooled results (Panels B and C), because our firm-specific time-series is short, and we are concerned about noisy estimation at the firm level.

Results for the firm-specific regressions in Panel A of Table 3 are consistent with the theory and the univariate results in Table 2. As predicted, current changes in working capital are negatively related to current cash flow from operations, and positively related to past and future cash flow from operations. The mean coefficient on current cash flows is −0.62, while the mean coefficients on past and future cash flows are 0.17 and 0.09, respectively. As discussed earlier (and detailed in Appendix B), the absolute magnitudes of the coefficients are less than the theoretical value of 1 because of the measurement error in the independent variables, and the downward bias is greater for the coefficients on past and future cash flows. Based on the cross-sectional distribution of the firm-specific coefficients, all of these means are highly statistically significant, with t-statistics ranging in absolute value from 10 to 57. Adjusted R^2s indicate that this specification provides reasonable explanatory power (mean $R^2 = 0.47$ and median $R^2 = 0.55$).

Results for industry-specific and pooled regressions in Panels B and C of Table 3 are consistent with the firm-specific results. The mean coefficients on current CFO are −0.51 for both the industry-specific regressions and the pooled regression compared with −0.62 for the firm-specific mean. The coefficients on past and future cash flows are also comparable in magnitude with the firm-specific results, on the magnitude of 0.15 to 0.19, and are reliably positive. The adjusted R^2s are lower, with an average of 0.34 for the industry specification and 0.29 for the pooled regression, perhaps because the implications of our firm-specific model are less descriptive in cross-sectional specifications. Summarizing, results for firm-specific, industry, and pooled specifications are similar. Based on better theoretical grounding and better empirical fit, we proceed with the firm-level specification.

We perform several tests of the robustness of our results. Recall that the cash flow variables (CFOs) are noisy estimates of the theoretical cash flow variables (cash flows related to accruals). Measurement error in the independent variables leads to both biased coefficient estimates and biased estimates of residuals. Our sensitivity checks rely on the observation that CFO is likely to be a good proxy for the theoretical cash flow variable when the firm is in a steady state, because in these circumstances, cash flows related to accruals are likely to be a fairly constant proportion of CFO.

We implement this observation on two dimensions. First, we rerun the regressions reported in Table 3, controlling for the effect of sales growth, either by including a growth term in the regression or by running the regression only on low-growth firms (defined as percentage sales growth between −5 and 5 percent). The tenor of the results remains the same for both specifications (results not included). In addition, controlling for growth does not appreciably affect the relations between accrual quality, firm characteristics, and persistence of earnings (discussed later). Second, we analyze the effect of cash flow volatility on accrual quality. If cash flow volatility is low, then we expect a relatively stable relation between total cash flows and cash flows related to accruals. We control for cash flow volatility either by including a cash flow volatility term in the regressions or by running the regression on low-volatility firms only (firms with volatility in the lowest quintile). Results remain qualitatively unchanged for both of these specifications.[5]

Finally, we investigate the effect of special items, which often contain long-term accruals, and therefore affect both our measure of short-term accruals and our measure of cash flow from operations (we assume that cash flow from operations is related only to short-term accruals). Specifically, sometimes Compustat includes restructuring charges in

---

[5] Because we believe that accrual quality is theoretically negatively related to cash flow volatility, the results after controlling for cash flow volatility are probably more correctly interpreted to indicate that the quality of accruals concept has construct validity beyond that captured by cash flow volatility. However, that does not imply that one *should* control for cash flow volatility in deriving accrual quality. Theoretically, high cash flow volatility causes low accrual quality because of large forecast errors in volatile environments, and the effect of this causal variable should not be excluded from the empirical construct. For this reason, we present the main results without a control for cash flow volatility. A similar argument applies to sales growth.

"other assets and liabilities" (item 307), contaminating our measure of short-term accruals. A portion of the Compustat cash flows from operations can also be related to restructuring charges, contaminating our measure of cash flows. When we replicate the results in Panel A of Table 3 for firms whose special items are less than 5 percent of earnings and 1 percent of assets, the tenor of the results is unchanged. When we exclude "other assets and liabilities" (item 307) from our measure of working capital accruals, the tenor of the results is also unchanged.

The Relation between Accrual Quality and Firm Characteristics

We use the standard deviation of the residuals as a firm-specific measure of accrual quality, where a higher standard deviation signifies lower quality.[6] This empirical measure of accrual quality can be utilized for a variety of purposes. For example, it can be used in market-based tests of the relation between stock prices and earnings. Another application is to adapt our measure of accrual quality to devise alternative tests of earnings management. Accruals manipulated by management in an opportunistic manner often behave empirically just like accruals that are the result of unavoidable errors in estimation. For example, from the point of view of the accounting system, booking a fictitious receivable and not collecting it looks similar to booking a regular receivable and not collecting it.[7]

In this section, we illustrate two applications of our measure of accrual quality. First, we explore the relation between our measure and selected firm characteristics. While we assume that the realization of individual estimation errors is random, the average *magnitude* of these errors for a particular firm is likely to be systematically related to the firm's characteristics such as the volatility of operations or the length of the operating cycle. For example, even with good skills and the best intentions, managers of firms in volatile industries are likely to make larger accrual estimation errors.

Mapping the relation between firm characteristics and estimation errors is important for research, practical, and pedagogical purposes. The reason is that a firm-specific regression estimate of the standard deviation of the residuals requires a lengthy time-series. In addition, the regression approach requires information about *future* cash flows, which reduces its usefulness in many settings (e.g., abnormal return strategies based on quality of earnings). Therefore, it is valuable to identify observable firm characteristics that act as instruments for the propensity to make estimation errors. We note that some firm characteristics we examine are likely to be correlated with the standard deviation of the residuals by construction (e.g., the standard deviation of accruals). This is not a concern, because our main interest is in identifying strong relations between the unobservable estimation errors and observable firm characteristics, regardless of the source of the relation.

Based on existing theory, results, and economic intuition, we expect that:

- **The longer the operating cycle, the lower accrual quality.**
  Longer operating cycles indicate more uncertainty, more estimation and errors of estimation, and therefore lower quality of accruals.
- **The smaller the firm, the lower accrual quality.**

---

[6] An alternative measure of accrual quality *at the firm-year level* is absolute value of the residual for that year. The tenor of the results is similar for this alternative specification, but the relations to firm characteristics and earnings persistence are weaker.

[7] Thus, as discussed earlier, the magnitude of estimation errors is also likely to be related to variables capturing managerial skill and managerial opportunism, e.g., meeting analysts' forecasts, auditor quality, stock offerings, and closeness to debt covenants.

We expect that large firms have more stable and predictable operations and, therefore, fewer and smaller estimation errors. In addition, large firms are likely to be more diversified and various portfolio effects across divisions and business activities reduce the relative effect of estimation errors.

- **The greater the magnitude of sales volatility, the lower accrual quality**
  Sales volatility indicates a volatile operating environment and the likelihood of greater use of approximations and estimation, with corresponding large errors of estimation and low accrual quality.
- **The greater the magnitude of cash flow volatility, the lower accrual quality**
  High standard deviation of cash flows is another measure of high uncertainty in the operating environment.
- **The greater the magnitude of accrual volatility, the lower accrual quality**
  Since our measure of accrual quality is derived as a residual from accruals, accrual volatility and accrual quality are at least partly related by construction.
- **The greater the magnitude of earnings volatility, the lower accrual quality**
  Earnings is the sum of cash flows and accruals. Since the volatility of both components is predicted to be negatively related to earnings quality, we expect that greater volatility in earnings signifies lower accrual quality.
- **The greater the frequency of reporting negative earnings, the lower accrual quality**
  Losses are indicative of severe negative shocks in the firm's operating environment. Accruals made in response to such shocks are likely to involve substantial estimation error (e.g., restructuring charges). Thus, losses are indicative of low accrual quality.
- **The greater the magnitude of accruals, the lower accrual quality**
  More accruals indicate more estimation and errors of estimation, and therefore lower quality of accruals.

Table 4 provides results for these hypothesized relations. Panel A presents descriptive statistics where some of the variables are first averaged on a firm basis to be consistent with the rest of the variables and our firm-level measure of accrual quality. The distribution of *Average operating cycle* (mean 141 days, standard deviation 62 days) indicates short operating cycles. (In unreported tests, we find that *all* firms in our sample have average operating cycles of less than one year.) Thus, our assumption that most working capital accruals reverse within one year seems reasonable for this sample.

Panel B of Table 4 presents Pearson correlations between our measure of accrual quality (*sresid*) and firm characteristics (results for Spearman correlations are similar). While all correlations have the predicted signs, and all are significant at the 0.0001 level, there is a large range of magnitudes. The correlations for length of the operating cycle (0.28) and standard deviation of sales (0.34) are comparatively low. The correlations for all remaining variables exceed 0.50, with highest correlations for the standard deviation of earnings (0.82), the standard deviation of accruals (0.75), and proportion of negative earnings (0.63). Such strong correlations suggest that these variables can be used as reliable instruments for accrual quality. The high correlation with the average level of working capital accruals (0.69) suggests a strong relation between *sresid* and the measure of accrual quality in Sloan (1996). We explore this relation in more detail later in the paper.

In Panel C of Table 4 we investigate whether parsimonious combinations of these firm characteristics capture accrual quality better than any individual variable. Our baseline specification regresses *sresid* on Std. Dev. Earn (the variable with the highest correlation in Panel B) and shows an adjusted R^2 of 0.67 (Model 1). Next, we decompose the volatility

TABLE 4
Descriptive Statistics and the Correlation between Quality of Working Capital Accruals (*sresid*) and Selected Firm Characteristics for 1,725 Firms between 1987 to 1999

Panel A: Descriptive Statistics

	Mean	*Standard Deviation*	*Lower Quartile*	*Median*	*Upper Quartile*
Standard deviation of the residuals (*sresid*)	0.028	0.025	0.011	0.020	0.037
Average operating cycle	141.08	61.807	90.229	131.505	184.81
Log (Total Assets)	5.50	2.16	3.88	5.44	7.06
Std. Dev. of Sales	0.215	0.205	0.097	0.166	0.273
Std. Dev. of CFO	0.060	0.040	0.031	0.051	0.078
Std. Dev. of ΔWC	0.056	0.038	0.025	0.045	0.076
Std. Dev. of Earn	0.050	0.044	0.021	0.037	0.063
Proportion of earnings that are negative for each firm	0.100	0.189	0	0	0.111
Average $\lvert\Delta$WC$\rvert$	0.048	0.033	0.023	0.040	0.067

Panel B: Pearson Correlations between the Standard Deviation of the Residuals (sresid) and Selected Firm Characteristics (p-values in parentheses)

Average Op. Cycle	*Log (Total Assets)*	*Std. Dev. Sales*	*Std. Dev. CFO*	*Std. Dev. ΔWC*	*Std. Dev. Earn*	*Prop. of Negative Earnings*	*Average $\lvert\Delta$WC$\rvert$*
0.28	−0.55	0.34	0.60	0.75	0.82	0.63	0.69
(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)

Panel C: Regressions Where the Dependent Variable is the Standard Deviation of the Residuals (sresid) and the Independent Variables are Firm Characteristics (n = 1,725)

		Intercept	*Std. Dev. Earn*	*Std. Dev. ΔWC*	*Std. Dev. CFO*	*Adj. R^2*
(1)	Coefficient	0.004	0.462			0.67
	(t-statistic)	(9.29)	(58.63)			
(2)	Coefficient	0.001		0.480	0.008	0.57
	(t-statistic)	(1.34)		(29.32)	(0.56)	
(3)	Coefficient	−0.002	0.317	0.257		0.76
	(t-statistic)	(−4.10)	(36.73)	(26.17)		

		Intercept	*Prop. Negative Earnings*	*Std. Dev. Sales*	*Average $\lvert\Delta$WC$\rvert$*	*Average Op. Cycle*	*Log (Total Assets)*	*Adj. R^2*
(4)	Coefficient	0.012	0.049	0.005	0.334	0.00	−0.001	0.61
	(t-statistic)	(6.03)	(21.62)	(2.63)	(21.79)	(0.07)	(−5.22)	

(*Continued on next page*)

TABLE 4 (Continued)

The standard deviation of the residuals (*sresid*) is calculated based on the residuals from the following firm-specific regressions:

$$\Delta WC_t = b_0 + b_1 CFO_{t-1} + b_2 CFO_t + b_3 CFO_{t+1} + \varepsilon_t$$

where:

Cash flow from operations (CFO) = item 308 from the Compustat Statement of Cash Flows;
Change in working capital (ΔWC) = ΔAR + ΔInventory − ΔAP − ΔTP + ΔOther Assets (net), where AR is accounts receivable, AP is accounts payable, and TP is taxes payable; and
Earnings before long-term accruals (Earn) = CFO + ΔWC.

All variables are scaled by average assets.

The standard deviations of Sales, Earn, CFO, and ΔWC are calculated at a firm level. Operating cycle is equal to 360/(Sales Average AR) + 360/(Cost of Goods Sold)/(Average Inventory). Proportion of earnings that are negative is calculated as the number of firm-years with negative earnings divided by the total number of firm-years for each firm. For this table, we use a log(Total Assets) specification to correct for right-tail skewness in assets.

of earnings into cash flow and accrual volatility, both of which are highly correlated with *sresid* in Panel B. Model (2) indicates that Std. Dev. ΔWC subsumes the explanatory power of Std. Dev. CFO. For a likely explanation, recall that *sresid* is *defined* as the variation of the component of accruals unrelated to cash flow realizations. Model (3) includes the volatility of working capital accruals and the volatility of earnings, with the R^2 increasing to 0.76. Finally, Model (4) regresses *sresid* on all remaining firm characteristics. The R^2 is 0.61, lower than the R^2 for the baseline regression with earnings volatility (Model 1). Summarizing, a simple practical way to gauge the quality of accruals is to assess the volatility of both earnings and accruals.

The Relation between Accrual Quality and Earnings Persistence

Earnings persistence is frequently discussed as a measure of earnings quality (e.g., Penman 2001, 623; Revsine et al. 2002, 245). The theory in this paper provides a natural link between our measure of accrual quality and earnings persistence. Specifically, an examination of Equation (3) reveals that, holding the time-series properties of realized cash flows constant, adding more accrual errors results in lower earnings persistence. Thus, we expect that firms with low accrual quality will also have low earnings persistence.

The results on the empirical relation between accrual quality and earnings persistence are presented in Panel A of Table 5. We report portfolio results to maintain comparability with earlier studies (e.g., Sloan 1996; Barth et al. 2001) and to examine for potential nonlinearities in the quality-persistence relation. The results have the same tenor in a regression specification (not tabulated). In Panel A, firm-years are first sorted into quintile portfolios based on the standard deviation of the firm-specific regression residuals (*sresid*) from estimating Expression (5). Within each portfolio, we regress future earnings on current earnings, and report the slope coefficient (called *Persistence*) and the adjusted R^2. We find a negative relation between standard deviation of the residuals and persistence; the slope coefficient declines from 0.94 to 0.55 between quintiles 1 and 5. This decline is monotonic but the relation is steeper in the tails. Since these are univariate regressions, the decline in adjusted R^2 from 0.83 to 0.28 parallels that of the slope coefficient. Thus, these results

TABLE 5
The Relative Information Content of Accrual Quality and Level of Accruals for Earnings Persistence for Firm-Years between 1987 to 1999

Panel A: Portfolios Based on the Magnitude of the Standard Deviation of the Residuals (sresid)

Portfolio	*Std. Dev. Resid (sresid)*	$\lvert\Delta WC_t\rvert$	*Persistence* (δ_1)	*Earnings Persistence Regression Adj.* R^2	*Number of Observations*
1	0.006	0.023	0.943	0.830	3,047
2	0.013	0.033	0.816	0.651	3,043
3	0.020	0.042	0.799	0.619	3,049
4	0.032	0.058	0.756	0.545	3,045
5	0.069	0.086	0.551	0.280	3,050

Panel B: Portfolios Based on the Absolute Value of Accruals ($\lvert\Delta WC_t\rvert$)

Portfolio	$\lvert\Delta WC_t\rvert$	*Std. Dev. Resid (sresid)*	*Persistence* (δ_1)	*Earnings Persistence Regression Adj.* R^2	*Number of Observations*
1	0.005	0.014	0.814	0.566	3,046
2	0.015	0.016	0.791	0.548	3,047
3	0.031	0.019	0.809	0.584	3,047
4	0.056	0.024	0.747	0.466	3,047
5	0.135	0.035	0.550	0.329	3,047

Accrual quality is measured as the standard deviation of working capital residuals from the following firm-specific regression:

$$\Delta WC_t = b_0 + b_1 CFO_{t-1} + b_2 CFO_t + b_3 CFO_{t+1} + \varepsilon_t$$

Level of accruals is measured as the absolute value of current change in working capital.
Earnings persistence (δ_1) is measured for each portfolio from the following regression:

$$Earn_{t+1} = \alpha + \delta_1 Earn_t + v_t$$

Cash flow from operations (CFO) = item 308 from the Compustat Statement of Cash Flows;
Change in working capital (ΔWC) = $\Delta AR + \Delta Inventory - \Delta AP - \Delta TP + \Delta Other\ Assets$ (net), where AR is accounts receivable, AP is accounts payable, and TP is taxes payable; and
Earnings before long-term accruals (Earn) = CFO + ΔWC.
All variables are scaled by average assets.

confirm the hypothesized positive relation between accrual quality and earnings persistence.[8]

Panel A of Table 5 also reports mean level of working capital accruals for each *sresid* portfolio. Similar to the results in Table 4, these portfolio results reveal a positive relation between standard deviation of the residuals and level of accruals. This relation is important because Sloan (1996) shows that level of accruals is less persistent than cash flows. Thus, at this point it is difficult to distinguish empirically between the effects of accrual quality and the level of accruals on earnings persistence. However, as mentioned earlier, there are *a priori* reasons to believe that our measure of accrual quality and the level of accruals are both proxies for the same aspect of the unobservable "true" accrual quality. Thus, to a large extent, the real issue is which of these two measures performs better empirically for the task at hand—in this case, explaining earnings persistence.

We pursue two approaches to distinguish between the explanatory power of *sresid* and level of accruals with respect to earnings persistence. In Panel B of Table 5, we form quintiles based on the absolute magnitude of accruals and measure persistence for each quintile.[9] Results confirm the expected negative relation between the level of accruals and earnings persistence; the slope coefficient declines from 0.81 to 0.55, and the adjusted R^2 from 0.57 to 0.33, between quintiles 1 and 5. However, these declines are less pronounced than those in Panel A. The relative decline in *Persistence* is 0.39 (0.55 for adjusted R^2) across the accrual quality quintiles vs. 0.26 (0.24 for adjusted R^2) across the level of accrual quintiles. In addition, the relation between level of accruals and persistence in Panel B is nonmonotonic, and essentially flat for low- to medium-level accruals. The overall impression from Panel B is that level of accruals indicates low earnings persistence only for extreme accrual realizations. Thus, the relation between accrual quality and earnings persistence is stronger than the relation between level of accruals and persistence, especially for firms with high accrual quality.

Our second approach is provided in Table 6, which reports portfolio results on the relation between *sresid* and earnings persistence, holding the level of accruals constant (Panel A); and the relation between the level of accruals and earnings persistence, holding *sresid* constant (Panel B). To hold the level of accruals constant (as in Panel A), we first sort the sample into decile portfolios based on level of accruals, and then within each level-of-accruals decile, we further sort the observations into five portfolios based on standard deviation of the residuals (subportfolios 1 to 5). We pool all ten subportfolios 1 together into one portfolio, all subportfolios 2 together into another portfolio, and so on. The result is five portfolios, with substantial variation of *sresid* but nearly the same level of accruals.

As shown in Panel A of Table 6, this two-pass portfolio construction controls quite well for level of accruals, with mean $|\Delta WC|$ ranging from 0.047 to 0.051 across portfolios 1 to 5. Despite the strong positive correlation between *sresid* and level of accruals at the firm level, the portfolios in Panel A exhibit large variation in *sresid*: the conditional range between mean *sresid* of portfolios 1 and 5 in Table 6 is 0.55 compared to the unconditional range of 0.63 in Table 5. This result indicates that the variation in *sresid* independent of

---

[8] Our measure of accrual quality has some "look-ahead" bias in explaining earnings persistence because good accrual quality means a positive relation between current accruals and future cash flows. An alternative specification, which uses residuals from regressing accruals on only past and present cash flows produces similar portfolio results. This robustness in results is probably explained by the relative stability of our measure of accrual quality over time.

[9] Our measure of accruals includes only changes in working capital, while the measure of accruals in Sloan (1996) includes both changes in working capital and depreciation. However, most of the variation in Sloan's measure is due to variation in accounts receivable and inventory (Sloan 1996, 297).

TABLE 6
The Incremental Information Content of Accrual Quality and Level of Accruals for Earnings Persistence for Firm-Years between 1987 to 1999

Panel A: Portfolios Based on the Magnitude of the Standard Deviation of the Residuals (sresid), Controlling for the Absolute Value of Accruals ($|\Delta WC_t|$)

Portfolio	*Std. Dev. Resid (sresid)*	$\lvert\Delta WC_t\rvert$	*Persistence* (δ_1)	*Earnings Persistence Regression Adj.* R^2	*Number of Observations*
1	0.008	0.047	0.913	0.783	3,044
2	0.015	0.047	0.819	0.642	3,049
3	0.022	0.047	0.742	0.542	3,046
4	0.032	0.049	0.666	0.426	3,052
5	0.063	0.051	0.602	0.328	3,043

Panel B: Portfolios Based on the Absolute Value of Accruals ($|\Delta WC_t|$), Controlling for the Standard Deviation of the Residuals (sresid)

Portfolio	$\lvert\Delta WC_t\rvert$	*Std. Dev. Resid (sresid)*	*Persistence* (δ_1)	*Earnings Persistence Regression Adj.* R^2	*Number of Observations*
1	0.006	0.027	0.788	0.502	3,042
2	0.020	0.028	0.762	0.513	3,049
3	0.036	0.028	0.743	0.438	3,050
4	0.059	0.028	0.707	0.457	3,049
5	0.120	0.029	0.574	0.408	3,044

Accrual quality is measured as the standard deviation of working capital residuals from the following firm-specific regression:

$$\Delta WC_t = b_0 + b_1 CFO_{t-1} + b_2 CFO_t + b_3 CFO_{t+1} + \varepsilon_t$$

Level of accruals is measured as the absolute value of current change in working capital.
Earnings persistence (δ_1) is measured for each portfolio from the following regression:

$$Earn_{t+1} = \alpha + \delta_1 Earn_t + v_t$$

Cash flow from operations (CFO) = item 308 from the Compustat Statement of Cash Flows;
Change in working capital (ΔWC) = ΔAR + ΔInventory − ΔAP − ΔTP + ΔOther Assets (net), where AR is accounts receivable, AP is accounts payable, and TP is taxes payable; and
Earnings before long-term accruals (Earn) = CFO + ΔWC.
All variables are scaled by average assets.

the level of accruals is substantial and empirically important. Further, the regression results reported in Table 6, Panel A reveal that the conditional variation in *sresid* is strongly related to earnings persistence. The slope coefficient varies from 0.91 in portfolio 1 to 0.60 in portfolio 5, and adjusted R^2 varies from 0.78 in portfolio 1 to 0.33 in portfolio 5. These results compare well to those from the unconditional specification in Table 5 (differences in *Persistence* of 0.31 vs. 0.39, and differences in R^2s of 0.45 vs. 0.54). In other words, the relation between accrual quality (*sresid*) and earnings persistence is largely preserved after controlling for the level of accruals.

Panel B in Table 6 contains the results for the relation between level of accruals and earnings persistence after controlling for accrual quality, using the same portfolio derivation method as in Panel A. First, we rank all observations into ten portfolios based on *sresid*, then within each decile portfolio we rank observations into five subportfolios on level of accruals, and finally we pool all subportfolios with the same rank together. At the end, we have five portfolios that preserve most of the variation in level of accruals, while holding *sresid* nearly constant across portfolios. Results reveal a slight decline in the variation of the slope coefficient across portfolios; the difference between portfolios 1 and 5 is 0.22 in Table 6 vs. 0.26 in Table 5. However, there is a substantial erosion in the range of R^2, the difference is 0.09 in Table 6 vs. 0.24 in Table 5. In addition, the differences in slope coefficients and R^2 across portfolios for Panel B in Table 6 are substantially smaller than their counterparts for *sresid* in Panel A.

Summarizing, the combined evidence of Tables 5 and 6 suggests that our measure of accrual quality captures variation in earnings persistence incremental to the variation explained by level of accruals. In fact, our results suggest that the relation between accrual quality and earnings persistence is stronger than that between level of accruals and persistence. However, relative to our measure of accrual quality, level of accruals is simpler and easier to use, and requires no extensive identification and prediction process. Thus, the comparison presented here is biased against the performance of the level of accruals, especially for settings that require *ex ante* measures of quality of earnings (e.g., real-time trading strategies).

V. CONCLUSION

This study suggests a new approach to assessing accrual and earnings quality, based on the intuition that accruals are temporary adjustments that resolve timing problems in the underlying cash flows at the cost of making assumptions and estimates. Precise estimates imply a good match between current accruals and past, present, and future cash flow realizations, while imprecise or erroneous estimates reduce the beneficial role of accruals. Following this intuition, we define accrual quality as the extent to which accruals map into cash flow realizations. In the empirical domain, we operationalize this notion of accrual quality as the standard deviation of the residuals from firm-specific regressions of working capital accruals on last-year, current, and one-year-ahead cash flow from operations.

One important feature of our approach is that the notion of estimation errors includes both intentional and unintentional errors. This distinction is important because most existing research assumes that accrual and earnings quality is only affected by management intent to manipulate, while such intent is unobservable, and likely idiosyncratic and sporadic. In contrast, our approach reveals that accrual quality is likely to be systematically related to observable and recurring firm characteristics like volatility of operations because higher volatility is associated with higher incidence of unavoidable estimation errors.

A theoretical limitation of our approach is that it captures *the extent* to which accruals map into the related cash flows but it provides little insight into the proper *timing* of these accruals with respect to the cash flows. For example, our approach cannot be used to decide whether to expense or capitalize R&D. Limitations in the empirical portion of our analysis include a restriction to working capital accruals and an assumption of no serial correlation in the estimation errors.

We illustrate the usefulness of our approach in two contexts. First, we document that observable firm characteristics can be used as instruments for accrual quality. Specifically, we find that earnings volatility and the volatility of accruals are good proxies for our measure of accrual and earnings quality. Second, we find a positive relation between accrual quality and earnings persistence, which remains after controlling for the level of accrual effect documented by Sloan (1996). This positive relation suggests there are important practical benefits from identifying and measuring accrual quality.

Our investigation of the interrelations between accrual quality, level of accruals, and earnings persistence also suggests a reconciliation of the findings of Dechow (1994) and Sloan (1996). Dechow (1994) finds that accruals improve earnings' ability to measure performance relative to cash flows. Sloan (1996) finds that the accrual portion is less persistent than the cash flow portion of earnings, which suggests that firms with high levels of accruals have low quality of earnings. Our reconciliation is based on the observation that a high level of accruals signifies both earnings that are a greater improvement over underlying cash flows, *and* low-quality earnings. The reason is that accruals are largest when the underlying cash flows have the most timing and mismatching problems, so more accruals signify greater improvement over the underlying cash flows. However, this benefit comes at the cost of incurring estimation errors, and there will be a positive correlation between levels of accruals and the magnitude of these estimation errors. Thus, everything else equal, large accruals signify low quality of earnings, and less persistent earnings.

APPENDIX A

Examples of Cash Flows and Accruals Types, and the Resulting Estimation Errors

The following table outlines the recording conventions when cash is received or paid before or after it is recognized in earnings:

	Cash Flow Occurs at t+1 *(after it is recognized in earnings)*	*Cash Flow Occurs at t−1* *(before it is recognized in earnings)*
	Case (1)	Case (3)
Recognized in earnings at t as revenue	Cash Inflow: Collection Accrual created and reversed: Asset e.g., Accounts Receivable	Cash Inflow: Deferred Revenue Accrual created and reversed: Liability e.g., Deferred Revenue
	Case (2)	Case (4)
Recognized in earnings at t as an expense	Cash Outflow: Payment Accrual created and reversed: Liability e.g., Warranty Liability	Cash Outflow: Deferred Cost Accrual created and reversed: Asset e.g., Inventory, Prepaid Rent

The examples below show the link between accruals and the related cash flows, and whether an estimation error will be detected by our model. We also include the notation introduced in Section II.

Case (1): Cash Receipt Follows Revenue Recognition
Record $100 of credit sales

t	Dr	Accounts Receivable ($A_{CFt+1/t}^{O}$)	100	
		Cr Sales Revenue		100

No error situation

t+1	Dr	Cash (CF_{t+1}^{t})	100	
		Cr Accounts Receivable ($A_{CFt+1/t}^{C}$)		100

Error situation: an unanticipated $30 is not collected

t+1	Dr	Bad debt write-off (ε_{t+1}^{t})	30	
	Dr	Cash (CF_{t+1}^{t})	70	
		Cr Accounts Receivable ($A_{CFt+1/t}^{C}$)		100

Cash increases by $70 at t+1 but accounts receivable at t increases by $100. Our model will detect an error of $30.

Case (2): Cash Payment Follows Expense Recognition
Estimate $100 for warranty expense

t	Dr	Warranty Expense	$100	
		Cr Warranty Liability ($A_{CFt+1/t}^{O}$)		$100

No error situation

t+1	Dr	Warranty Liability ($A_{CFt+1/t}^{C}$)	100	
		Cr Cash (CF_{t+1}^{t})		100

Error situation: Warranty work amounts to $135

t+1	Dr	Warranty Liability ($A_{CFt+1/t}^{C}$)	100	
	Dr	Warranty expense (ε_{t+1}^{t})	35	
		Cr Cash (CF_{t+1}^{t})		135

Warranty liability increases by $100 at t, and is not equal to the cash flow of $135 at t+1. Our model will detect a $35 estimation error.

Case (3): Cash Receipt Precedes Revenue Recognition
Receive $100 for a prepaid magazine subscription

t−1	Dr	Cash (CF_{t-1}^{t})	100	
		Cr Deferred Revenue ($A_{CFt-1/t}^{O}$)		100

No error situation

t	Dr	Deferred Revenue ($A_{CFt-1/t}^{C}$)	100	
		Cr Sales Revenue		100

Error situation: Customer demands a refund of $10 at time t

t	Dr	Deferred Revenue ($A_{CFt-1/t}^{C}$)	100	
		Cr Cash		10
		Cr Sales Revenue		90

Our model will not detect an estimation error because the deferred revenue accrual and cash flow at t–1 are equal in magnitude, and the deferred revenue reversal at t is equal to the cash flow received at t–1. Note that accrual earnings contains no "error" since only \$90 is recognized in earnings, while cash basis revenue recognition would overstate revenue (by \$10) at t–1 and understate revenue by the same amount at t. This type of error will affect our empirical version of the model because we cannot identify specific cash flows relating to accruals, and our measure of cash flows at t will include the \$10. This effect is similar to that induced by including cash expenses CF_t^t, in our measure of cash flows CFO_t (see Appendix B).

Case (4): Cash Payment Precedes Expense Recognition
Pay \$100 for inventory

t–1	Dr	Inventory ($A_{CFt-1/t}^{O}$)	100	
		Cr Cash (CF_{t-1}^{t})		100

No error situation

t	Dr	Cost of goods sold	100	
		Cr Inventory ($A_{CFt-1/t}^{C}$)		100

Error situation: Write down of \$20

t	Dr	Write down	20	
	Dr	Cost of goods sold	80	
		Cr Inventory ($A_{CFt-1/t}^{C}$)		100

Our model will not detect an estimation error for the write-down because the inventory accrual and cash flow at inception are equal, and the inventory reversal at t is equal to the cash flow paid at t–1. Our model, therefore, provides no insights into the *timing* of expense recognition.

APPENDIX B

Simulations Examining the Effect of Mismeasuring Cash Flows in Our Regressions

Our theoretical model of accruals listed in equation (4) is the following:

$$A_t = CF_{t-1}^{t} - (CF_t^{t+1} + CF_t^{t-1}) + CF_{t+1}^{t} + \varepsilon_{t+1}^{t} - \varepsilon_t^{t-1}. \tag{4}$$

Our empirical approximation of this model is:

$$\Delta WC_t = b_0 + b_1{*}CFO_{t-1} + b_2{*}CFO_t + b_3{*}CFO_{t+1} + \varepsilon_t. \tag{5}$$

The cash flow terms in Equation (5) capture the constructs in Equation (4) with measurement error, where the measurement error is given by bolded components in parentheses reported below:

$$\begin{aligned} CFO_{t-1} &= CF_{t-1}^{t} && + (\mathbf{CF_{t-1}^{t-1}} + \mathbf{CF_{t-1}^{t-2}}) \\ CFO_t &= (CF_t^{t+1} + CF_t^{t-1}) && + (\mathbf{CF_t^{t}}) \\ CFO_{t+1} &= CF_{t+1}^{t} && + (\mathbf{CF_{t+1}^{t+1}} + \mathbf{CF_{t+1}^{t+2}}) \end{aligned}$$

For CFO_{t-1}, the noise consists of cash flow received and recognized in earnings in the same period (CFO_{t-1}^{t-1}) plus cash payments/collections (e.g., receivables) for transactions recognized in earnings at $t-2$. For CFO_t the noise stems from cash flows received and recognized in earnings in the same period. For CFO_{t+1} the noise stems from cash flows received and recognized in earnings in the same period (CFO_{t+1}^{t+1}) plus deferred cash flows to be recognized in the following period (CFO_{t+1}^{t+2}). Thus, with respect to current accruals, the noise component for CFO_t is smaller than the noise component for CFO_{t-1} and CFO_{t+1}.

We use simulations to examine the effect from this type of noise on the coefficients and R^2 in Equation (5). In these simulations there is no estimation error. All credit sales are collected, all deferred revenues are recognized the following period, and there are no write-downs of inventory. Thus, any change in the R^2 and the coefficients is entirely attributable to mismeasurement of the desired cash flows.

Our model firm generates sales in three ways:

(1) Credit sales are made at time t and collected in t+1 (creates accounts receivable)
(2) Cash sales are made at time t
(3) Deferred revenue is recognized at time t for cash received at t-1 (creates deferred revenue)

We use a random number generator to generate each type of sales for 200 periods. Sales randomly range from 0 to 200. The firm starts business with an initial endowment of inventory and shareholders' equity. The inventory policy is to replace in period t inventory sold in period $t-1$, and there are no price changes. For simplicity, we assume that the inventory required for cash sales at t is purchased at t so that cash from operations and earnings are the same for cash sales (i.e., no accruals are made for cash sales). The initial endowment is set large enough such that the firm never has stock-outs of inventory. In this firm:

Working capital = Accounts Receivable + Inventory − Deferred Revenue;
Accruals = ΔAccounts Receivable + ΔInventory − ΔDeferred Revenue;
Cash from operations in period t (CFO_t) = Cash from credit sales made at $t-1$ (CF_t^{t-1}) + Cash for cash sales made at t (CF_t^t) + Deferred Revenue (CF_t^{t+1}) − Cash outlays for inventory (CF_t^t); and
Earnings = Cash from operations + Accruals = Sales − Cost of Goods Sold.

Exhibit 1 provides regression results for a benchmark case, where all cash flows are measured correctly, as in Equation (4). Since we have perfect knowledge of the related cash flows, and there is no estimation error, we expect that $b_1 = 1$, $b_2 = -1$, $b_3 = 1$, and the R^2 is 100 percent. The simulation confirms this intuition.

Exhibit 2 documents the effect of using cash flow from operations (CFO, as defined above) instead of the theoretically correct cash flow variables. Since the cash flow variables are measured with error, we expect R^2 and the coefficients to decrease in absolute magnitude. In addition, we expect that the larger measurement error for CFO_{t-1} and CFO_{t+1} will force b_1 and b_3 more toward zero as compared to b_2. These predictions are confirmed by the results in Exhibit 2.

EXHIBIT 1

Regression Results of Simulations Where There Is No Mismeasurement of Cash Flows at t−1, t, or t+1

$$\Delta WC_t = b_0 + b_1(CF_{t-1}^{t}) + b_2(CF_t^{-1} + CF_t^{+1}) + b_3(CF_{t+1}^{t}) + \varepsilon_t$$

	Intercept	b_1	b_2	b_3	*Adjusted* R^2
Coefficient	0.00	1.00	−1.00	1.00	1.00
(p-value)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	

EXHIBIT 2

Regression Results of Simulations Where Cash Flows from Operations (CFO) for t−1, t, and t+1 Are Used in the Regressions

$$\Delta WC_t = b_0 + b_1{*}CFO_{t-1} + b_2{*}CFO_t + b_3{*}CFO_{t+1} + \varepsilon_t$$

	Intercept	b_1	b_2	b_3	*Adjusted* R^2
Coefficient	−0.75	0.49	−0.94	0.49	0.74
(p-value)	(0.623)	(< 0.001)	(< 0.001)	(< 0.001)	

Summarizing, our simulations confirm that using cash flow from operations (as in Equation (5)) instead of the theoretically prescribed cash flow variables (as in Equation (4)) results in reduced R^2 and regression coefficients biased toward zero, with less pronounced effect for the coefficient of current CFO.

REFERENCES

Anthony, J., and K. Petroni. 1997. Accounting estimation disclosures and firm valuation in the property-casualty insurance industry. *Journal of Accounting, Auditing, and Finance* 12 (Summer): 257–281.

Barth, M. E., D. P. Cram, and K. K. Nelson. 2001. Accruals and the prediction of future cash flows. *The Accounting Review* 76 (January): 27–58.

Beaver, W. H., and M. McNichols. 1998. The characteristics and valuation of loss reserves of property casualty insurers. *Review of Accounting Studies*: 73–95.

Collins, D. W., and P. Hribar. 2002. Errors in estimating accruals: Implications for empirical research. *Journal of Accounting Research* 40: 105–134.

Dechow, P. M. 1994. Accounting earnings and cash flows as measures of firm performance: The role of accounting accruals. *Journal of Accounting and Economics* 18 (July): 3–42.

———, S. P. Kothari, and R. L. Watts. 1998. The relation between earnings and cash flows. *Journal of Accounting and Economics* 25 (May): 133–168.

Financial Accounting Standards Board (FASB). 1978. *Objectives of Financial Reporting by Business Enterprises.* Statement of Accounting Concepts No. 1. Norwalk, CT: FASB.

———. 1987. *The Statement of Cash Flows.* Statement of Financial Accounting Standards No. 95. Norwalk, CT: FASB.

Finger, C. A. 1994. The ability of earnings to predict future earnings and cash flow. *Journal of Accounting Research* 32: 210–223.

Healy, P. M., and J. M. Wahlen. 1999. A review of the earnings management literature and its implications for standard setting. *Accounting Horizons* 13 (December): 365–383.

Liu, J., D. Nissim, and J. Thomas. 2002. Equity valuation using multiples. *Journal of Accounting Research* 40 (March): 135–172.

Palepu, K. G., P. M. Healy, and V. L. Bernard. 2000. *Business Analysis and Valuation*. Cincinnati, OH: South-Western College Publishing,

Penman, S. 2001. *Financial Statement Analysis and Security Valuation*. New York, NY: McGraw-Hill, Irwin.

Petroni, K. R. 1992. Optimistic reporting in the property-casualty insurance industry. *Journal of Accounting and Economics* 15 (December): 485–509.

Revsine, L., D. W. Collins, and W. B. Johnson. 2002. *Financial Reporting and Analysis*. Upper Saddle River, NJ: Prentice Hall.

Sloan, R. G. 1996. Do stock prices fully reflect information in accruals and cash flows about future earnings? *The Accounting Review* 71 (July): 289–315.

THE ACCOUNTING REVIEW
Vol. 77
Supplement 2002
pp. 61–69

DISCUSSION OF The Quality of Accruals and Earnings: The Role of Accrual Estimation Errors

Maureen F. McNichols
Stanford University

ABSTRACT: Dechow and Dichev (2002) model earnings quality as the magnitude of estimation errors in accruals, and provide empirical estimates of this construct based on the relation between accruals and cash flows. I characterize the innovation and limitations in this approach, and provide empirical evidence of measurement error in their empirical specification. I also adapt their model to assess the specification of the Jones' (1991) model and document that this model provides estimates of discretionary accruals that are significantly associated with cash flows, which are likely to be substantially nondiscretionary. I conclude with suggestions for future research on earnings quality and earnings management.

Keywords: *earnings quality; accruals; cash flows; earnings management.*

I. INTRODUCTION

The paper by Dechow and Dichev (2002) (hereafter DD) proposes a measure of the quality of accruals and earnings based on the extent to which accruals map into cash flow realizations in contemporaneous and adjacent time periods. Underpinning their measure of earnings quality is the notion that accruals are estimates of future cash flow realizations, so that the quality of accruals and earnings is an inverse function of the precision of these estimates. Economic and structural factors can cause variation in the precision of accruals estimates (across accounts for a given firm, for a given firm over time, and across firms), regardless of the presence or absence of managerial interventions in the reporting process. In addition, managerial expertise will affect the precision of estimates, even if other factors that affect precision are held constant. As a result, a less-than-perfect mapping between accruals and cash flows in adjacent periods can reflect firms that report honestly but face uncertain economic environments, firms whose managers are less expert

---

I gratefully acknowledge the helpful comments of participants at *The Accounting Review* Conference on Quality of Earnings. I also especially appreciate the comments of Stephen Ryan and Katherine Schipper, and the able research assistance of Yulin Long.

at estimation, and or firms whose managers intervene in the process to manipulate accruals. My discussion aims to shed light on the contribution of DD's paper by characterizing its innovation and limitations, and by putting it in the context of related literature. My discussion concludes with suggestions for future research on earnings quality.

II. DECHOW AND DICHEV'S DEFINITION OF EARNINGS QUALITY

The accounting literature includes several definitions of earnings quality. Some define earnings as high quality if earnings are persistent, an attribute based solely on the time-series properties of earnings. Some define earnings as high quality if earnings accurately represent the economic implications of underlying transactions and events. Some, and this includes DD, define earnings quality in terms of the relation between accruals and cash flows. DD's definition does not distinguish among the various factors that influence this relation—the uncertainty in the firm's environment, the ability of management, the extent to which accruals are manipulated—but some approaches to defining earnings quality do.[1]

DD characterize the linkage between current accruals and cash flows in the immediately adjacent periods. Recognizing that accruals may arise following some cash flows and in anticipation of others, they develop a model that reflects estimation error in anticipated cash flows.[2] They characterize aggregate accruals as the sum of opening and closing deferral and accrual entries. Specifically:

$$A_t = CF_{t-1}{}^{t} - (CF_t{}^{t+1} + CF_t{}^{t-1}) + CF_{t+1}{}^{t} + \varepsilon_{t+1}{}^{t} - \varepsilon_t{}^{t-1} \qquad \text{(DD4)}$$

where:

A_t = current accruals recognized in period t;
$CF_t{}^{s}$ = cash from operations realized in period t and recognized in period s; and
$\varepsilon_t{}^{s}$ = estimation error associated with accruals recognized in period s and cash flows realized in period t.

The first and fourth terms reflect recognition of cash flows realized in t−1 and to be realized in t+1. The second term defers recognition of cash flows realized in t that are to be recognized in t+1, and the third term reflects cash flows realized in t that were recognized in t−1. The fifth and sixth terms reflect the estimation error in period t's opening accrual that will be realized in t+1 and the closing error for period t−1 realized in period t.

Estimation error is defined as the difference between the amount accrued and the amount realized. Therefore period t's earnings include the opening error that will be realized in t+1 when the related cash flows are realized and the reversing error from the cash flows realized in period t that differ from the amount accrued in period t−1. DD define the quality of accruals and earnings as the magnitude of these errors.

Cash flow realizations in periods prior to t−1 and subsequent to t+1 are assumed to be beyond the horizon reflected in current accruals. That is, the approach they take requires an estimated coefficient for every period in which there is a pertinent accrual. In a model of, say, deferred taxes or depreciation, this is a long series of leads/lags indeed, and the model is not operational. The approach *requires* that the key element of accruals/earnings quality is in the current accruals. This is a significant assumption that limits the applicability of their approach to firms with operations that are shorter-term in nature.

---

[1] See for example, Jonas and Blanchet (2000) for discussion of dimensions of financial reporting quality, Penman and Zhang (2002) for discussion of the role of accounting choices and estimates on the sustainability of earnings, and the literature on earnings management, (e.g., Healy [1985], Moses [1987], McNichols and Wilson [1988], Schipper [1989], Bernard and Skinner [1996], Demski and Frimor [1999], Healy and Wahlen [1999] for discussion of management's incentives to intervene in the financial reporting process).

[2] Their model does not consider however, subsequent revision of a prior accrual, such as an inventory impairment.

The DD definition of earnings quality differs from that of Comiskey and Mulford (2000) who define earnings as high quality if the contemporaneous cash flows are greater (less) than the recognized revenues or gains (expenses or losses), and low quality if the associated cash flows are less than (greater than) the recognized revenues or gains (expenses or losses). This view considers a firm with significant unearned revenue and deferred expenses to have high-quality earnings, and a firm with significant accrued revenue and prepaid expenses to have low-quality earnings. In contrast, DD define earnings to be of equal quality for firms with high vs. low realizations of the sum of the error terms in Equation (4) if the variance of the sum of the errors for the firms is equal. Their notion of earnings quality assumes a symmetric loss function, that investors are indifferent between estimation errors that overstate and understate future cash flow realizations by an equal amount.

The estimation errors are assumed to be independent of each other and of the cash flow realizations. The discussion in the paper suggests this assumption is plausible, but this arises from the focus on the behavior of total accruals. DD do not separately consider how total accruals might be affected by the behavior of discretionary accruals. Prior literature suggests that estimation errors caused by management discretion are not likely to be independent of each other and of the cash flow realizations, and therefore suggests that DD's analytical model and predictions may not apply in a context of great interest, where management intervenes in the financial reporting process to alter perceptions of the firm's underlying performance.

For example, Demski and Frimor (1999) develop a model in which discreteness in the payoffs to the agent causes the magnitude of manipulation to be nonlinear in what might be called pre-managed earnings. Healy (1985) suggests discretionary and nondiscretionary accruals will be uncorrelated when pre-managed earnings are less than the earnings target and management cannot manipulate upward sufficiently to achieve positive results, but will be negatively correlated when earnings are greater than the target. McCulloch and Black (2002) show that, if management alters accruals in response to fluctuations in nondiscretionary accruals to achieve a smoother income series, then nondiscretionary and discretionary accruals will be negatively correlated. Furthermore, they argue that discretion in accruals can induce correlation in errors across years, as managers exercise income-increasing (-decreasing) discretion in low- (high-) earnings years and reverse this discretion in subsequent high- (low-) earnings years.

Empirical evidence also documents that discretionary accrual estimates are correlated with earnings performance. Dechow et al. (1995) and Kasznik (1999) find that firms with higher (lower) earnings exhibit significantly positive (negative) discretionary accruals, suggesting earnings management varies with earnings or that the Jones (1991) model used to estimate nondiscretionary accruals is misspecified. Beaver et al. (2002) find that property and casualty insurance firms' loss reserve errors are significantly more negative for firms with large positive earnings than for firms in other regions of the earnings distribution, suggesting firms in this industry with large positive earnings use discretion to lower earnings. Thus, the theory and evidence suggests incorporating management's incentives to exercise discretion over accruals in the model will result in different implications than those of the present model.

III. EMPIRICAL IMPLEMENTATION

DD characterize the "empirical version" of Equation (4) as:

$$\Delta WC_t = b_0 + b_1 CFO_{t-1} + b_2 CFO_t + b_3 CFO_{t+1} + \varepsilon_t$$

and measure earnings quality as the standard deviation of the residual. This section discusses three issues raised by the empirical implementation of the DD earnings quality measure: the standard deviation of the residual as a proxy for earnings quality, the difference between reported cash from operations and the cash flows related to period t, and the implications of common firm transactions for the residual term.

First, the standard deviation of the residual, in theory, is the estimated square root of the variance of the dependent variable conditional on the explanatory variables. This measure therefore reflects absolute variation in the residual rather than variation relative to the variation in accruals. Holding the R^2 of the model constant, firms with more variable accruals will have a higher standard deviation of the residuals, so firms with greater underlying volatility in earnings are classified as having lower quality earnings. In this sense, the empirical results in the paper relating DD's empirical proxy to the absolute magnitude of accruals, and to the standard deviations of sales, cash flows, accruals, and earnings, are likely to be induced mechanically, in that variability in the residual is increasing in variability in accruals, which in turn is correlated with variability in sales, cash flows, and earnings.

The inverse empirical measure of quality in the paper is the standard deviation of residual accruals, but, and given the mechanics of OLS regression, this measure will be larger for firms with large absolute accruals holding relative estimation error constant. The results would be more readily interpretable if the *relative* variance of the errors were shown to be associated with proxies for environments in which it is difficult to forecast. In other words, for firms where the R^2 in DD's empirical equation is high (low), one would expect, other things equal, an environment where forecasting ability is great (small).

A second issue with the empirical specification, which DD note, is that the cash from operations data available to researchers includes cash flows realized in a period that are recognized in multiple periods, in contrast to the theoretical specification, which includes only cash flows recognized in period t but realized in periods t−1 or t+1, and cash flows realized in t but recognized in periods t−1 and t+1. Their empirical specification has error in the independent variables, leading to biased coefficients. They analyze this issue in Appendix B and their simulations indicate that the effect of the measurement error is to bias downward the coefficients on all three cash from operations variables in their equation.

However, in addition to the amounts included in cash from operations that are outside the model, a number of other items can affect the analysis. For example, to the extent sample firms participate in mergers and acquisitions, or in divestitures, the accrual data in one period may not correspond to the cash flow data in another. Collins and Hribar (2001) provide evidence that these transactions are fairly common during their sample period 1988–1997, with over a third of their sample engaging in mergers and or discontinued operations. These transactions can cause current accruals to be based on a different entity than future cash flows. A related issue arises for rapidly growing firms. The cash from operations in t+1 may be greater than amounts accrued in t due to growing sales. In this case, cash from operations in period t+1 contains additional measurement error, as it includes the cash flow effects of sales for period t+1 as well as for period t, and is therefore likely biased downward further.

IV. EARNINGS QUALITY AND DISCRETIONARY ACCRUALS

The purpose of this section is to link the DD analysis of earnings quality to the literature on discretionary accruals, and in particular, the Jones (1991) model. Jones' intent was to *separate* discretionary accruals (DA) from nondiscretionary accruals (NDA). DD's intent is to assess accruals as a whole—they do not attempt to separate management-induced effects (the analog of DA) from all other effects. Linking the approach taken by DD with that taken by Jones (1991) has the potential to strengthen both approaches, and to calibrate the errors associated with Jones' measure of discretionary accruals and DD's measure of earnings quality.

As discussed in McNichols (2000), there are many reasons to suspect that the estimated discretionary accruals from the Jones model reflect nondiscretionary forces rather than pure discretion. In particular, the Jones model assumes accruals react to the current change in sales, but that lagged and future changes are not relevant. Counter to this assumption are the assumptions in Bernard and Stober (1989) and Dechow et al. (1998) that accruals do not fully adjust to a contemporaneous sales shock; rather, adjustment occurs over succeeding periods. Furthermore, anticipation of future sales growth is likely to influence management's estimates, as reflected in accruals. Consistent with this, McNichols (2001) documents that analysts' long-term earnings growth forecasts have significant explanatory power for residuals estimated using the Jones model, suggesting that growth is a significant correlated omitted variable in this model.

The estimation results in DD suggest that including cash flows in the Jones model might reduce the extent to which the model omits variables that are correlated with sample firms' economic fundamentals. Relatedly, measurement error in DD's estimation may preclude them from controlling for the fundamental factors influencing accruals. Therefore, including sales in the DD model provides a useful specification check on the magnitude of measurement error in their cash flow variables. To assess this, I compare the incremental explanatory power of the independent variables in each of these models for the other, to provide evidence on the validity of each specification.

To test for a relation between discretionary accrual estimates from the Jones (1991) model and cash flows, I examine a sample of all nonfinancial firms on the Compustat annual research files during 1988–1998.[3] This sample period permits me to use SFAS No. 95 (FASB 1987) statement of cash from operations data to estimate accruals, rather than a balance sheet approach. Given the findings of Collins and Hribar (2000), I exclude companies with mergers, acquisitions, or discontinued operations, as indicated by Compustat footnote 1.[4]

To examine whether cash flows have explanatory power for accruals after controlling for the change in current period sales and the level of plant and equipment, and vice versa, I estimate the following equations:

$$\Delta WC_t = b_0 + b_1\, CFO_{t-1} + b_2\, CFO_t + b_3\, CFO_{t+1} + \varepsilon_t \tag{1}$$

$$\Delta WC_t = b_0 + b_1\, \Delta Sales_t + b_2\, PPE_t + \varepsilon_t \tag{2}$$

$$\Delta WC_t = b_0 + b_1\, CFO_{t-1} + b_2\, CFO_t + b_3\, CFO_{t+1} + b_4\, \Delta Sales_t + b_5\, PPE_t + \varepsilon_t \tag{3}$$

---

[3] The sample includes firms with SIC codes from 2000–3999.

[4] Specifically, all firm-year observations coded by Compustat as experiencing an accounting change, merger or acquisition, or discontinued operations or a fiscal year change that affected sales were excluded.

where ΔWC_t is defined as in DD to include the change in accounts receivable (item 302), the change in inventory (item 303), the change in accounts payable (item 304), the change in taxes payable (item 305) and change in other assets (item 307); CFO_t is cash from operations from the SFAS No. 95, *Statement of Cash Flows*, (item 308); and the $\Delta Sales_t$. (item 12) and PPE_t (item 7) are the change in sales and the level of property, plant, and equipment, respectively.[5] Equation (1) is DD's empirical model, as characterized by their Equation (5) and Equation (2) is the Jones model. Equation (3) includes the explanatory variables from Equations (1) and (2), to assess whether included variables in either of Equations (1) or (2) have explanatory power for the other equation.

The estimation results for the above equations are in Table 1. The findings confirm that, consistent with DD, accruals are significantly positively associated with prior year, current year, and subsequent year cash from operations; the R^2 is .20. The coefficient estimates are 0.0543, −0.2214, and 0.1312, with associated t-statistics of 23.33, −60.18, and 41.72, respectively. Consistent with DD, the strongest association is with current cash from operations.

---

[5] Jones (1991) includes depreciation in her measure of accruals, but to allow for consistency with DD, the measure of accruals adopted here excludes depreciation.

TABLE 1
Estimation Results from Regression of Changes in Working Capital on Cash Flows, Change in Sales and Plant and Equipment

This table presents estimation results for the sample of 15,015 firm-year observations with available data from the Compustat Annual Industrials and Research file, with SIC codes 2000–3999 for years 1988–1998. Observations with merger and acquisitions, accounting changes, fiscal year changes, and discontinued operations are excluded.

$$\Delta WC_t = b_0 + b_1\, CFO_{t-1} + b_2\, CFO_t + b_3\, CFO_{t+1} + \varepsilon_t \qquad (1)$$

$$\Delta WC_t = b_0 + b_1\, \Delta Sales_t + b_2\, PPE_t + \varepsilon_t \qquad (2)$$

$$\Delta WC_t = b_0 + b_1\, CFO_{t-1} + b_2\, CFO_t + b_3\, CFO_{t+1} + b_4\, \Delta Sales_t + b_5\, PPE_t + \varepsilon_t \qquad (3)$$

Estimation	*Intercept*	CFO_{t-1}	CFO_t	CFO_{t+1}	$\Delta Sales_t$	PPE_t	*Adjusted* R^2	*F-value*	*n*
1	0.0154	0.0543	−0.2214	0.1312			0.2011	1260.9	15015
	27.57	23.33	−60.18	41.72					
2	0.00951				0.0802	−0.0052	0.0730	592.26	15015
	0.57				33.74	−5.57			
3	0.0096	0.0645	−0.2409	0.1251	0.0966	−0.0027	0.3009	1293.38	15015
	12.18	29.41	−69.31	42.48	45.93	−3.32			

Variable definitions:

ΔWC_t = changes in working capital accounts as disclosed on the statement of cash from operations, measured as the increase in accounts receivable (Compustat data item #302) plus the increase in inventory (#303) plus the decrease in accounts payable and accrued liabilities (#304) plus decrease in taxes accrued (#305) plus the increase (decrease) in other assets (liabilities) (#307), deflated by beginning total assets;

CFO_t = cash from operations in year t (Compustat data item #308);

ΔSALES = change in sales (data item #12) deflated by beginning total assets (data item #6); and

PPE = property, plant, and equipment deflated by beginning total assets (data item #6).

The next set of estimation results is for the Jones model; these results indicate significant explanatory power, with an adjusted R^2 of .07 and F-statistic of 592.3, but substantially less than the estimation based on DD's model. Accruals and the change in sales are significantly positively associated, as indicated by a coefficient of 0.0802 and t-statistic of 33.74.

The third set of estimation results combines cash flow variables from DD's model with Jones model variables. The findings indicate that both models are likely significantly misspecified, in that both the cash flow variables and the change in sales remain highly significant; the R^2 increases to .30. The coefficient on the change in sales is 0.0966 with a t-statistic of 45.93.

The findings in Table 1 suggest misspecification in both the Jones and DD models. Specifically, the estimation results for Equation (3) indicate that the residual from Equation (1) is significantly correlated with the change in sales, consistent with cash from operations being a noisy proxy for the cash flows recognized in accruals. The findings in Table 1 also indicate that the residual from Equation (2), the Jones model, is significantly associated with lagged, current, and future cash flows. Because these variables most likely reflect fundamentals to a greater extent than discretion, the findings indicate that estimates of discretionary accruals based on the Jones model likely include a substantial nondiscretionary component.

The findings indicate that researchers should consider the implications of both the DD and Jones models to develop more powerful approaches to the estimation of earnings quality and the role of management discretion in influencing earnings quality. As McNichols and Wilson (1988) discuss, to the extent a proxy for discretionary accruals contains measurement error that is correlated with the partitioning variables in a study's research design, inferences are affected. This holds for studies of earnings quality as well. Taken as a whole, the findings suggests that further modeling of the relation between accruals and cash flows could yield substantial improvements in our ability to understand the factors that influence earnings quality. This should also yield substantial improvements in our ability to test for management's exercise of discretion over accruals.

V. CONCLUDING REMARKS

In my view, the paper by Dechow and Dichev (2002) makes two contributions. First, it provides a characterization of the relation between accruals and cash flows that captures an important element of earnings quality—estimation error in accruals. This paper extends the modeling in Dechow (1994), Dechow et al. (1998), and Barth et al. (2001) by incorporating the role of estimation error, and by drawing an explicit link to earnings quality.

The second contribution is to operationalize this characterization empirically and provide some evidence on its validity. Prior literature has provided evidence on the contemporaneous relation between accruals and cash flows (Rayburn 1986; Bowen et al. 1986; McNichols and Wilson 1988; Bernard and Stober 1989; Dechow 1994; Dechow et al. 1998). Prior literature has also provided evidence on the relation between accruals and future cash flows (Finger 1994; Dechow et al. 1998; Barth et al. 2001). The present paper corroborates many of the findings in these earlier studies. In addition, it suggests a rationale for an empirical specification relating accruals to cash flows from the prior, contemporaneous, and subsequent periods.

The approach taken in this paper suggests several future research directions. The first direction is to enrich the modeling by specifying the process generating cash flows, the information available to management, management's estimation task, and the consequences of these estimates. One can then incorporate management's incentives to report objective vs. biased accruals, and determine the implications of these enrichments for the relation between accruals and cash flows. The second direction is to focus on specific accruals rather than aggregate accruals. The complexity associated with modeling the estimation errors in aggregate accruals is daunting, and the construct validity associated with a proxy based on aggregate accruals seems low. A focus on specific accruals can permit a more complete characterization of the relation between accruals and cash flows, and can potentially result in a better understanding of the role played by estimation error. For example, studies in the property and casualty insurance area (e.g., Petroni 1992; Petroni et al. 2000; Beaver and McNichols 1998, 2001) focus directly on loss reserve errors, which reflect the change in management's estimates of claim loss reserves over time. The potential to develop models of specific accruals such as restructuring reserves, warranty liabilities, and specific components of current accruals is largely untapped. Such development can potentially allow better understanding of the forces shaping management's choices and their relation to the measurement error in earnings.

REFERENCES

Barth, M., D. Cram, and K. Nelson. 2001. Accruals and the prediction of future cash flows. *The Accounting Review* 76 (1): 27–58.

Beaver, W. H., and M. F. McNichols. 1998. The characteristics and valuation of loss reserves of property casualty insurers. *Review of Accounting Studies* 3 (1-2): 73–95.

———, and ———. 2001. Do stock prices of property casualty insurers fully reflect information about earnings, accruals, cash flows and development? *Review of Accounting Studies* 6 (2-3): 197–220.

———, ———, and K. Nelson. 2002. Management of the loss reserve accrual and the distribution of earnings in the property-casualty insurance industry. Working paper, Stanford University.

Bernard, V. L., and T. Stober. 1989. The nature and amount of information in cash flows and accruals. *The Accounting Review* 64 (4): 624–652.

———, and D. Skinner. 1996. What motivates managers' choice of discretionary accruals? *Journal of Accounting and Economics* 22 (1/3): 313–325.

Bowen, R., D. Burgstahler, and L. A. Daley. 1986. Evidence on the relationships between earnings and various measures of cash flow. *The Accounting Review* 61: 713–725.

Collins, D. W., and P. Hribar. 2001. Errors in estimating accruals: Implications for empirical research. Working paper, The University of Iowa.

Comiskey, E., and C. W. Mulford. 2000. *Guide to Financial Reporting and Analysis*. New York, NY: John Wiley & Sons.

Dechow, P. M. 1994. Accounting earnings and cash flows as measures of firm performance: The role of accounting accruals. *Journal of Accounting and Economics* 18 (1): 3–42.

———, R. G. Sloan, and A. P. Sweeney. 1995. Detecting earnings management. *The Accounting Review* 70 (2): 193–225.

———, S. P Kothari, and R. L. Watts. 1998. The relation between earnings and cash flows. *Journal of Accounting and Economics* 25 (2): 133–168.

———, and I. D. Dichev. 2002. The quality of accruals and earnings: The role of accrual estimation errors. *The Accounting Review* (Supplement): 35–59.

Demski, J., and H. Frimor. 1999. Performance measure garbling under renegotiation in multi-period agencies. *Journal of Accounting Research* 37 (Supplement): 187–214.

Financial Accounting Standards Board (FASB). 1987. *Statement of Cash Flows*. Statement of Financial Accounting Standards No. 95. Norwalk, CT: FASB.

Finger, C. 1994. The ability of earnings to predict future earnings and cash flows. *Journal of Accounting Research* 32 (2): 210–223.

Healy, P. M. 1985. The effect of bonus schemes on accounting decision. *Journal of Accounting and Economics* 7 (1-3): 85–107.

Jonas, G., and J. Blanchet. 2000. Assessing quality of financial reporting. *Accounting Horizons* 14 (3): 353–363.

Jones, J. J. 1991. Earnings management during import relief investigation. *Journal of Accounting Research* 29 (2): 193–228.

Kasznik, R. 1999. On the association between voluntary disclosure and earnings management. *Journal of Accounting Research* 37 (1): 57–81.

McCulloch, B. W., and E. L. Black. 2002. Theoretical and empirical relations among components of accruals under earnings management in a multi-period setting. Working paper, New Zealand Treasury and Brigham Young University.

McNichols, M., and G. P. Wilson. 1988. Evidence of earnings management from the provision for bad debts. *Journal of Accounting Research* 26 (Supplement): 1–31.

———. 2000. Research design issues in earnings management studies. *Journal of Accounting and Public Policy* 19 (4-5): 313–345.

Moses, O. D. 1987. Income smoothing and incentives: Empirical tests using accounting changes. *The Accounting Review* 62 (2): 358–377.

Penman, S., and X. Zhang. 2002. Accounting conservatism, the quality of earnings and stock returns. *The Accounting Review* 77 (2): 237–264.

Petroni, K. 1992. Optimistic reporting in the property-casualty insurance industry. *Journal of Accounting and Economics* 15 (4): 485–508.

———, S. G. Ryan, and J. M. Wahlen. 2000. Discretionary and non-discretionary revisions of loss reserves by property-casualty insurers: Differential implications for future profitability, risk, and market value. *Review of Accounting Studies* 5 (2): 95–125.

Rayburn, J. 1986. The association of operating cash flow and accruals with security returns. *Journal of Accounting Research* 24 (Supplement): 112–133.

Schipper, K. 1989. Commentary on earnings management. *Accounting Horizons* 3 (4): 91–102.

THE ACCOUNTING REVIEW
Vol. 77
Supplement 2002
pp. 71–105

The Relation between Auditors' Fees for Nonaudit Services and Earnings Management

Richard M. Frankel
Massachusetts Institute of Technology
Marilyn F. Johnson
Michigan State University
Karen K. Nelson
Stanford University

ABSTRACT: This paper examines whether auditor fees are associated with earnings management and the market reaction to the disclosure of auditor fees. Using data collected from proxy statements, we present evidence that nonaudit fees are positively associated with small earnings surprises and the magnitude of discretionary accruals, while audit fees are negatively associated with these earnings management indicators. We also find evidence of a negative association between nonaudit fees and share values on the date the fees were disclosed, although the effect is small in economic terms.

Keywords: *auditor independence; auditor fees; earnings management; discretionary accruals.*

Data Availability: *The data used in this study are from the public sources identified in the text.*

I. INTRODUCTION

This paper provides empirical evidence on the relation between auditor fees and earnings management, and between auditor fees and share values on the date the fees are disclosed. Our tests are motivated by concerns about the effect of auditors' provision of nonaudit services on the credibility of financial reports. These concerns were

---

We thank Bill Beaver, Walt Blacconiere, Larry Brown, Joe Carcello, S. P. Kothari, Tom Linsmeier, Roger Martin, Maureen McNichols, Andrew Sbaraglia, Joe Weber, Scott Whisenant, two anonymous reviewers, and seminar participants at *The Accounting Review* Conference on Quality of Earnings at Emory University; University of British Columbia; University of California, Berkeley; Carnegie Mellon University; Georgia State University; Michigan State University; Northwestern University; Pennsylvania State University; and the 2001 Stanford Summer Accounting Camp for helpful discussions and comments. Professor Nelson acknowledges the support of the Financial Research Initiative at Stanford University Graduate School of Business. We thank First Call for providing the analyst forecast data. Waqas Nazir, Nora Richardson, and Xiaoyan Zheng provided valuable data collection assistance.

a primary reason that the Securities and Exchange Commission (SEC) issued revised auditor independence rules in November 2000. The rules require firms to disclose, in proxy statements filed on or after February 5, 2001, the amounts of all audit fees and nonaudit fees paid to the auditor of their financial statements.

The provision of nonaudit services can strengthen the auditor's economic bond with the client, thereby increasing the auditor's incentive to acquiesce to client pressure, including pressure to allow earnings management (e.g., Simunic 1984; Beck et al. 1988a; Beeler and Hunton 2001). However, the provision of nonaudit services can also increase the auditor's investment in reputational capital, which the auditor is not likely to jeopardize to satisfy the demands of any one client (Arruñada 1999). We test these competing hypotheses by examining management's deliberate intervention in the earnings process in the form of earnings management. We rely on conventional measures of earnings management to proxy for these unobservable interventions. Specifically, we examine the propensity to just meet or beat earnings benchmarks and the magnitude of discretionary accruals to assess whether the provision of nonaudit services is associated with earnings management.

Disclosure of auditors' fees for nonaudit services can inform investors about financial reporting quality, including earnings management. If investors believe the provision of nonaudit services compromises auditor objectivity and they require compensation for a perceived decline in the credibility of firms' earnings reports, then they will bid down the share values of firms disclosing unexpectedly high nonaudit fees (Simunic 1984; Teoh and Wong 1993). However, if investors believe the provision of nonaudit services increases the audit firm's investment in reputational capital, thereby increasing the likelihood of credible reporting, then share prices will increase for firms disclosing unexpectedly large nonaudit fees. We test these competing hypotheses by examining the association between nonaudit fees and share values on the date the fees are disclosed in proxy statements. However, we cannot rule out the possibility that investors infer information unrelated to the credibility of financial reports from the disclosures of nonaudit fees, such as information about management quality and operational difficulties.

Although recent concerns about auditor independence have focused on the provision of nonaudit services to audit clients, it is possible that audit fees create similar bonding or reputational incentives. Thus, we also investigate the association between earnings management, share prices, and fees for audit services. However, our tests for both types of fees are only suggestive of the incentives arising from auditor-client contracting; we assume the incentive effects of audit and nonaudit fees are captured in the observable indicators of earnings management and share prices, but we do not provide direct evidence of auditor behavior. In this regard, our archival evidence complements experimental (Beeler and Hunton 2001) and survey (Nelson et al. 2002) studies of auditor incentives.

Our analysis is based on data collected from 3,074 proxy statements filed with the SEC between February 5, 2001, and June 15, 2001. We develop three complementary specifications of auditor fees. The first specification, the ratio of nonaudit fees to total fees (FEERATIO), is suggested by the SEC and prior research. The second specification disaggregates auditor fees into nonaudit (RANKNON) and audit (RANKAUD) components, both of which we measure as percentile ranks for each client of a given auditor. This specification captures the client's financial importance to the auditor, and allows for a separate test of the incentive effects of audit and nonaudit fees. The third specification is the percentile rank of total auditor fees for each client of a given auditor (RANKTOT), and is consistent with claims that audit and nonaudit fees create similar incentive effects.

We find a significant positive association between the purchase of nonaudit services, proxied by FEERATIO and RANKNON, and the likelihood of reporting a small earnings surprise, the magnitude of absolute discretionary accruals, and the magnitude of income-increasing and income-decreasing discretionary accruals. These results are robust to alternative measures of the earnings management variables, including discretionary total accruals, discretionary working capital accruals, and performance-matched discretionary accruals. Additional evidence from auditor-specific regressions indicates that the results are not driven by any particular auditor. However, the relation between nonaudit fees and small earnings surprises is insignificant for the larger firms in the sample. There is also no association between nonaudit fees and the likelihood of reporting a small increase in earnings.[1] These results provide some evidence that firms purchasing nonaudit services manage earnings to a greater extent than other firms.

In contrast to these results, we find a significant negative association between audit fees, proxied by RANKAUD, and the earnings management indicators, including small increases in earnings. However, this effect is less pronounced for the largest firms in the sample. There is no association between the proxy for total fees, RANKTOT, and the indicators of earnings management. Taken together, these results suggest that audit and nonaudit fees create different incentive effects, contrary to arguments in prior research (e.g., Hansen and Watts 1997; Reynolds and Francis 2000), and that combining the fees into a single measure masks these effects.

Finally, we find a significant negative association between share prices and the disclosure of higher than expected nonaudit fees. However, the association is small in economic terms, and is insignificant when abnormal returns are measured over longer event windows. There is no evidence of an association between share prices and the disclosure of audit fees.

The remainder of the paper is organized as follows. Section II discusses the SEC's new rules on auditor independence and develops our hypotheses. Section III describes the research design. Section IV presents the empirical results. Section V summarizes and concludes.

II. BACKGROUND AND HYPOTHESIS DEVELOPMENT

SEC Rule on Auditor Independence

In November 2000, the SEC issued Final Rule S7-13-00, *Revision of the Commission's Auditor Independence Requirements* (hereafter the Rule). The Rule defines independence as "a mental state of objectivity and lack of bias" and adopts a two-pronged test to determine if an auditor is independent. The first prong refers to the auditor's mental state, i.e., independence "in fact." However, mental states are not observable, and thus the second prong provides for an assessment of independence "in appearance." The SEC will not consider an auditor to be independent with respect to a particular client "if a reasonable investor, with knowledge of all relevant facts and circumstances, would conclude that the auditor is not capable of exercising objective and impartial judgment" (SEC 2000, Section I).

The Rule requires companies to disclose, in proxy statements filed on or after February 5, 2001, information about fees billed by the auditor for the most recent year. The disclosures are intended to provide information useful to investors in evaluating whether nonaudit fees have impaired the auditor's independence (SEC 2000, Section II.c.5). Firms are required to disclose:

---

[1] We do not test a third earnings benchmark considered in prior literature, i.e., small positive earnings, because there is no evidence in our sample, or for the Compustat comparison sample in 2000, of a significant discontinuity in the distribution of earnings immediately around zero that would be considered evidence of earnings management to avoid losses.

(1) under the caption "Audit Fees," the aggregate fees billed for the audit of the annual financial statements and for reviews of the quarterly financial statements;
(2) under the caption "Financial Information Systems Design and Implementation Fees," the aggregate fees billed for professional services rendered for (i) directly or indirectly operating or supervising the company's information system or local area network, or (ii) designing or implementing hardware or software that aggregates financial statement source data or data significant to the financial statements; and
(3) under the caption "All Other Fees," the aggregate fees billed for all services rendered by the auditor other than those described in (1) and (2).

The Rule narrowly defines audit fees to include only those fees for services necessary in the audit of 10-Ks and the review of 10-Qs. Fees for other services associated with the attest function, such as pension plan audits, regulatory audits, due diligence procedures for mergers and acquisitions, and internal audit services, are included in other fees. Similarly, fees for auditing the tax provision and tax accrual are included in audit fees, while fees for other tax services, such as preparation and filing of tax returns and tax planning advice, are included in other fees. Thus, the category "All Other Fees" is likely to be more heterogeneous with respect to the nature of fees included than either of the other categories.

Hypothesis Development

Auditor Fees and Earnings Management

Simunic (1984) and Beck et al. (1988a) model the joint demand for audit and nonaudit services, and show that when the same auditor provides both services and the auditor retains a portion of the cost savings from "knowledge spillovers," the auditor will be bonded economically to the client. An auditor concerned about the possible loss of nonaudit fee revenue is less likely to object to management's accounting choices if being dismissed as auditor makes it less likely the auditor will be hired to provide nonaudit services.

Empirical evidence of economic bonding due to the provision of nonaudit services is mixed. Simunic (1984) documents that the purchase of nonaudit services from the auditor is associated with a significant increase in audit fees, and interprets this result as consistent with the hypothesis that the joint provision of audit and nonaudit services creates knowledge spillovers that could lead to economic bonding. However, Palmrose (1986) and Davis et al. (1993) argue that this result is due to underlying economic factors that drive the demand for both nonaudit and audit services. Abdel-khalik (1990) finds no interdependencies between audit and nonaudit fees. Beck et al. (1988b) examine economic bonding in the context of auditor tenure, and find that auditor tenure is longer and less variable when the client purchases a high level of nonaudit services from the auditor. Finally, Parkash and Venable (1993) and Firth (1997), using U.S. and U.K. data, respectively, find evidence that firms with high agency costs purchase relatively smaller amounts of nonaudit services from their auditors to avoid jeopardizing independence or the appearance of independence.

In contrast to the agency literature that characterizes auditor bias as deliberate, the behavioral literature suggests that psychological heuristics unconsciously lead auditors to bias judgments. For example, Beeler and Hunton (2001) argue that when the realization of future economic rents depends on maintaining an audit relationship, auditors are favorably predisposed toward their clients. They present experimental evidence that audit partners are more likely to assume the client is a going concern and to revise downward their initial estimates of the number of hours budgeted for the audit when there are contingent economic rents arising from the provision of nonaudit services.

Although the cause of auditor bias differs in the agency and behavioral literatures (i.e., intentional distortion vs. cognitive bias), the hypothesized effect is the same; auditors are more likely to acquiesce to client pressure, including pressure to allow earnings management, when the provision of nonaudit services generates economic rents. This argument assumes that the audit partner benefits from the provision of nonaudit services to his or her client. Consistent with this assumption, Beeler and Hunton (2001) find a high correlation (0.72, p-value < 0.01) between nonaudit revenue generated by an audit partner's client and the partner's personal income for a sample of 73 audit partners surveyed from four of the Big 5 audit firms.

The competing hypothesis predicts that auditor objectivity will not be compromised by providing nonaudit services to audit clients. Arruñada (1999) develops a model showing that the provision of nonaudit services to audit clients increases the audit firm's investment in reputational capital. This investment constrains auditor bias because the gains from acquiescing to any one client's demands are outweighed by the reputational losses that would be imposed by other clients who value a reputation for independence. Dopuch et al. (2001) demonstrate analytically that the increase in reputational capital resulting from the provision of nonaudit services increases auditor independence provided the probability of misstatement risk is low. However, Trompeter (1994) argues that the audit firm's investment in reputational capital may not constrain bias at the individual partner level where compensation depends on retaining a handful of clients.

Auditors also face substantial legal sanctions for failed audits. Aggregate legal claims against the then Big 6 auditing firms exceeded \$30 billion at the end of 1992, although subsequent legal developments have reduced auditors' exposure to legal liability for securities fraud (Coffee 2001). However, there is little evidence to suggest that the provision of nonaudit services is a significant factor in damage awards, or that liability insurers consider the provision of nonaudit services to be a risk factor in pricing liability insurance contracts (Antle et al. 1997).

Because of the competing theoretical arguments and the mixed empirical evidence, we test the following nondirectional hypothesis, stated in null form:

> **H1a:** The provision of nonaudit services to audit clients is not associated with earnings management.

Prior literature discusses similar incentive effects for audit fees. For example, DeAngelo (1981) and Magee and Tseng (1990) argue that economic rents associated with audit fees create an economic bond between auditor and client, and therefore incentives for auditors to permit earnings management; however, high audit fees also increase the firm's investment in reputational capital, providing a disincentive for auditors to permit earnings management. Thus, we also test the following nondirectional hypothesis, stated in null form:

> **H1b:** The provision of audit services is not associated with earnings management.

Audit and nonaudit fees may create similar incentive effects, as argued by Hansen and Watts (1997) and Reynolds and Francis (2000). However, we are not aware of any evidence supporting or refuting this claim. Thus, our empirical models allow for separate estimation of the association between audit fees, nonaudit fees, and proxies for earnings management.

Market Reaction to the Disclosure of Auditor Fees

Incentive arguments also suggest an association between share prices and the disclosure of auditor fees. If investors believe the provision of nonaudit services compromises auditor objectivity and reduces the probability of credible reporting, then they will bid down share values of firms disclosing unexpectedly large nonaudit fees (Simunic 1984; Teoh and Wong 1993). Consistent with a negative market reaction to revelations about the credibility of earnings reports, Firth (1990) observes a decline in market values of client firms at the release of U.K. Department of Trade reports criticizing the work of their auditors. Similarly, Feroz et al. (1991) and Dechow et al. (1996) document a decline in stock prices around the announcements of SEC enforcement actions for manipulating earnings.[2] However, if investors believe the provision of nonaudit services to audit clients increases the audit firm's investment in reputational capital, thereby increasing the probability of credible reporting, then share prices will increase for firms disclosing unexpectedly large nonaudit fees. Similar arguments apply to the economic bonding and reputational capital effects of audit fees. Thus, we test the following nondirectional hypotheses, stated in null form:

H2a: There is no share price reaction to the disclosure of nonaudit fees.

H2b: There is no share price reaction to the disclosure of audit fees.

III. RESEARCH DESIGN

Sample Description

Table 1, Panel A summarizes the sample selection procedure. Our initial sample consists of 4,701 proxy statements on the SEC's EDGAR database with a filing date between February 5, 2001 and June 15, 2001. From this total, we exclude 1,124 proxy statements filed by financial institutions (SIC codes 6000–6999) because of the unique procedures required to estimate discretionary accruals for these firms. We also exclude 51 proxy statements that do not relate to an annual meeting of shareholders because the SEC does not require disclosures of auditor fees in these proxy statements. Finally, we exclude 71 proxy statements indicating that the firm changed auditors during the year because the change is likely to affect the amount of fees billed. Of the remaining 3,455 observations, 218 do not disclose the required components of auditor fees.[3] We match the firms with fee data to Compustat, resulting in a final sample of 3,074 firms.

Panel B of Table 1 details the frequency of proxy filings by month. Firms must file prior to their annual shareholders' meetings, which typically occur three to four months after the fiscal year end. Most (about 83 percent) of our sample firms filed their proxy statements in March or April, corresponding to a December or January fiscal year end. Panel C of Table 1 reveals that the industry composition of our sample is similar to that of the firms on Compustat in 2000.

Table 2, Panel A reports statistics for the three categories of fees required under the SEC rule. The mean (median) ratio of audit fees to total fees is 0.51 (0.49), with a minimum of 0.02 and a maximum of 1.00. Only 8 percent of sample firms hired their auditors to provide financial information systems design services (IS). Although the mean (median) ratio of IS fees to total fees is 0.02 (0.00), untabulated findings indicate this ratio is 0.28 (0.17) for firms incurring these fees. The mean (median) ratio of other fees to total fees is 0.47 (0.48), with 96 percent of the sample reporting fees in this category. Finally, the mean (median) ratio of nonaudit fees (i.e., the sum of IS and other fees) to total fees is 0.49 (0.51).

---

[2] We do not expect the stock price response to the disclosure of the fee data to be as extreme as in these studies because we do not focus on firms identified by the SEC as violating Generally Accepted Accounting Principles.

[3] The noncompliance rate for our sample is thus 6 percent (i.e., 218/3,455), well below the rate of 17 percent reported by Abbott et al. (2001). However, their sample includes only proxy statements filed between February 5, 2001 and March 16, 2001. The noncompliance rate for our sample during this period is 13 percent, while the noncompliance rate for firms filing after March 16, 2001 is 5 percent. Abbott et al. (2001) report that firms failing to disclose auditor fees are significantly smaller and more likely to have a non-Big 5 auditor.

Approximately one-third of the sample (1,052 firms) voluntarily discloses information on the composition of other fees. We code these fees into four categories: (1) audit-related, e.g., audits of employee benefit plans, regulatory audits, and preparation of registration statements and other SEC filings; (2) tax, e.g., preparation and filing of tax forms and tax-related consulting; (3) combined audit-related and tax, consisting of categories (1) and (2); and (4) other advisory, e.g., general consulting services, and information technology consulting for systems not associated with the financial statements.[4]

---

[4] The majority (60 percent) of firms disclosing other advisory fees are clients of Ernst & Young, and these firms generally include tax fees with other advisory fees.

TABLE 1
Sample Description

Panel A: Selection Procedure for Sample of Firms with Auditor Fee Disclosures

Selection Criteria	*Observations*
Definitive proxies filed between 2/5/01 and 6/15/01	4,701
Less: Financial institutions (SIC codes 6000–6999)	(1,124)
Proxies not relating to annual meeting of shareholders	(51)
Change in auditor during fiscal year 2000	(71)
	3,455
Missing auditor fee data	(218)
Observations with auditor fee data available	3,237
Less: Observations not available on Compustat	(163)
Sample Observations	3,074

Panel B: Distribution of Observations by Month of Proxy Filing

Month	*n*	*%*
February	80	2.60
March	771	25.08
April	1,769	57.55
May	366	11.91
June	88	2.86
Total	3,074	100.00

(*Continued on next page*)

TABLE 1 (Continued)

Panel C: Distribution of Observations by Industry

	Sample		*2000 Compustat*
Industry Description	*n*	*%*	*%*
Agriculture	9	0.29	0.43
Mining and construction	66	2.15	2.74
Food	45	1.46	2.30
Textiles and printing/publishing	148	4.81	4.88
Chemicals	71	2.31	2.47
Pharmaceuticals	244	7.94	5.88
Extractive	124	4.03	4.33
Durable manufacturers	733	23.85	23.77
Transportation	228	7.42	7.09
Utilities	109	3.55	3.15
Retail	353	11.48	10.97
Services	344	11.19	11.78
Computers	600	19.52	20.22
Total	3,074	100.00	100.00

Industry membership is determined by SIC code as follows: agriculture (0100–0999), mining and construction (1000–1999, excluding 1300–1399), food (2000–2111), textiles and printing/publishing (2200–2799), chemicals (2800–2824, 2840–2899), pharmaceuticals (2830–2836), extractive (1300–1399, 2900–2999), durable manufacturers (3000–3999, excluding 3570–3579 and 3670–3679), transportation (4000–4899), utilities (4900–4999), retail (5000–5999), services (7000–8999, excluding 7370–7379), and computers (3570–3579, 3670–3679, 7370–7379).

Panel B of Table 2 reports descriptive statistics of the components of other fees. Audit-related fees are disclosed by 21 percent of the sample. Tax and other advisory fees are disclosed by 13 and 11 percent of the sample, respectively, while 6 percent of the sample reports combined audit-related and tax fees. The voluntarily disclosed fees are a substantial portion of other fees, ranging from a mean (median) of 46 (45) percent for other advisory fees to 95 (100) percent for combined audit-related and tax fees. Untabulated regression results confirm that the likelihood of voluntary disclosure increases with the magnitude of other fees relative to total fees. Moreover, conditional on the ratio of other fees to total fees, firms audited by Ernst & Young are more likely to disclose components of other fees, while firms audited by the other Big 5 auditors and by non-Big 5 auditors are less likely to disclose.

Table 3 summarizes the fee data for each of the Big 5 audit firms, and for all other audit firms combined. Among the Big 5, total fees billed to the sample firms range from $1.77 billion for PricewaterhouseCoopers to just over $700 million for KPMG. Total fees billed to sample firms by non-Big 5 auditors are $64 million. The Big 5 and non-Big 5 also differ in the composition of total fees. Audit fees comprise no more than one-third of

TABLE 2
Descriptive Statistics of Auditor Fees Disclosed in Proxy Statements Filed between February 5, 2001 and June 15, 2001
(n = 3,074)

Panel A: Mandatory Disclosures of Fee Data

Variable	*Mean*	*Standard Deviation*	*First Quartile*	*Median*	*Third Quartile*	*Minimum*	*Maximum*	*% > 0*
Audit	511	1,495	104	191	425	5	48,000	100
Audit/Total	0.51	0.24	0.31	0.49	0.70	0.02	1.00	
IS	209	1,939	0	0	0	0	46,800	8
IS/Total	0.02	0.11	0.00	0.00	0.00	0.00	0.95	
Other	1,050	3,784	59	211	674	0	77,000	96
Other/Total	0.47	0.24	0.28	0.48	0.66	0.00	0.94	
Nonaudit	1,258	4,797	60	221	722	0	79,700	96
Nonaudit/Total	0.49	0.24	0.30	0.51	0.69	0.00	0.98	
Total	1,769	5,945	191	439	1,143	10	103,600	

Panel B: Voluntary Disclosures of Components of Other Fees, Divided by Other Fees

Variable	*Mean*	*Standard Deviation*	*First Quartile*	*Median*	*Third Quartile*	*Minimum*	*Maximum*	*% Discl.*
Audit-Related	0.57	0.31	0.31	0.55	0.87	0.01	1.00	21
Tax	0.64	0.34	0.31	0.66	1.00	0.02	1.00	13
Combined Audit-Related and Tax	0.95	0.16	1.00	1.00	1.00	0.10	1.00	6
Other Advisory	0.46	0.29	0.20	0.45	0.69	0.01	1.00	11

(Continued on next page)

TABLE 2 (Continued)

Audit = aggregate fees billed for professional services rendered for the audit of the annual financial statements and the reviews of the quarterly financial statements;
IS = aggregate fees billed for financial information systems design and implementation;
Other = aggregate fees billed for all services rendered other than the services covered by Audit and IS;
Nonaudit = sum of IS and Other; and
Total = total fees billed.

The components of Other fees in Panel B are as follows: Audit-Related, e.g., audits of employee benefit plans, regulatory audits, and preparation of registration statements and other SEC filings; Tax, e.g., preparation and filing of tax forms, and tax-related consulting; Combined Audit-Related and Tax, consisting of fees from both of the previous categories; and Other Advisory, e.g., general consulting services, and information technology consulting for systems not associated with the financial statements.

All levels are in thousands of dollars.

% Discl. indicates the percentage of observations disclosing the components of Other fees.

TABLE 3
Descriptive Statistics of Auditor Fees Disclosed in Proxy Statements Filed between February 5, 2001 and June 15, 2001, By Auditor
(n = 3,074)

			Composition of Total Fees				Median Ratio of Fees, by Component, to Total Fees			
Auditor[a]	*n*	*Total Fees*	*Audit*	*IS*	*Other*	*Nonaudit*	*Audit*	*IS*	*Other*	*Nonaudit*
AA	605	1,110,570	0.33	0.12	0.54	0.67	0.45	0.00	0.50	0.55
DT	376	723,945	0.31	0.11	0.58	0.69	0.50	0.00	0.47	0.50
EY	682	1,058,961	0.28	0.06	0.66	0.72	0.46	0.00	0.53	0.54
KPMG	436	708,523	0.29	0.15	0.56	0.71	0.50	0.00	0.47	0.50
PWC	681	1,771,831	0.25	0.14	0.61	0.75	0.42	0.00	0.54	0.58
All Other	294	64,348	0.58	0.00	0.42	0.42	0.73	0.00	0.27	0.30

[a] The auditors are Arthur Andersen (AA), Deloitte & Touche (DT), Ernst & Young (EY), KPMG (KPMG), and PricewaterhouseCoopers (PWC). There are 91 separate auditors included in the category of "All Other," including Grant Thornton (84 observations), BDO Seidman (51 observations), and McGladrey & Pullen (17 observations). No other auditor exceeds 15 observations.

Total Fees = total fees billed for Audit, IS, and Other, in thousands of dollars;
Audit = aggregate fees billed for professional services rendered for the audit of the annual financial statements and the reviews of the quarterly financial statements;
IS = aggregate fees billed for financial information systems design and implementation;
Other = aggregate fees billed for all services rendered other than the services covered by Audit and IS; and
Nonaudit = sum of IS and Other.

total fees billed for any of the Big 5 auditors, compared to 58 percent of total fees billed by non-Big 5 auditors. The composition of fees also varies across the Big 5 auditors. Ernst & Young earns only 6 percent of its revenue from IS services, but bills more for other services (66 percent) than any other auditor. Nonaudit fees range from 67 percent of total fee revenue for Arthur Andersen to 75 percent for PricewaterhouseCoopers.

The remaining columns of Table 3 report, by auditor, the median of the client-specific ratios of fees billed in each category to total fees. The median ratio of nonaudit fees to total fees is 0.30 for the non-Big 5 auditors, while for Big 5 auditors the ratio is at least 0.50. Thus, at least half of the Big 5's audit clients pay nonaudit fees in excess of audit fees. However, the median ratio of nonaudit fees to total fees is lower than the proportion of total fee revenue derived from nonaudit fees. Thus, a relatively small number of clients account for a disproportionate amount of each Big 5 auditor's nonaudit fee revenue.

The findings in Tables 2 and 3 differ markedly from 1999 data considered by the SEC in formulating the auditor independence rule (SEC 2000). The 1999 data suggest that nonaudit fees comprise 10 percent of total fees, with almost 75 percent of audit clients purchasing no nonaudit services. In contrast, our findings indicate that nonaudit fees comprise approximately 70 percent of total fee revenue (Table 3), with 96 percent of audit clients purchasing nonaudit services (Table 2, Panel A). Although the categorization of fees used to compile the 1999 data may not be the same as that required by the SEC rule, this comparison illustrates that the actual magnitude and frequency of nonaudit services is considerably greater than previously reported.

Model Specification

Auditor Fee Variables

We use the auditor fee disclosures to develop three specifications of auditor fees. The first specification, the ratio of nonaudit fees to total fees (FEERATIO), is consistent both with the SEC's position that the proportion of fees for nonaudit services is useful to investors in assessing auditor independence (SEC 2000, Section III.c.5), and with prior research (e.g., Scheiner 1984; Glezen and Millar 1985; Beck et al. 1988b; Parkash and Venable 1993; Barkess and Simnett 1994; Firth 1997). However, this ratio is invariant to the scale of the fees, and thus does not capture the financial importance of the client to the auditor. In addition, cross-sectional variation in the ratio is driven by the levels of both nonaudit fees and audit fees.

To capture the client's financial importance to the auditor, the second specification is the percentile rank, by auditor, of the amount of nonaudit fees (RANKNON) and audit fees (RANKAUD) disclosed by each firm.[5] Firms in the highest (lowest) percentile receive a rank of 100 (1). We use percentile ranks rather than scaling by total audit firm revenue to mitigate measurement error (Johnston 1984). Disaggregating auditor fees into components also permits us to test the separate incentive effects of nonaudit and audit fees. Including both variables in the specification mitigates correlated omitted variables bias as audit and nonaudit fees are positively correlated (e.g., Simunic 1984; Craswell et al. 1995).

The final specification is the percentile rank, by auditor, of the amount of total fees disclosed by each firm (RANKTOT). This measure assumes similar incentive effects of nonaudit and audit fees, and, like FEERATIO, is affected by cross-sectional variation in the levels of both nonaudit and audit fees.

---

[5] DeFond et al. (2002) and Ashbaugh et al. (2002) use a log transformation of audit and nonaudit fees. The rank-transformed variables and the log-transformed variables are highly correlated (Spearman correlations > 0.90). Our results are robust to this alternative specification of the fee variables.

Estimation Equations

We use two complementary approaches to identify earnings management. Prior research (e.g., Burgstahler and Dichev 1997; Burgstahler and Eames 1998; Degeorge et al. 1999) documents a higher than expected frequency of slightly positive earnings surprises, earnings changes, and earnings levels, consistent with earnings management to meet simple benchmarks. The SEC and the Public Oversight Board (POB) expressed particular concern that the pressure to meet analysts' expectations and to project a smooth earnings path creates pressure on auditors to permit their clients to manage earnings to meet those objectives (SEC 2000; POB 2000).

Using First Call data, we calculate earnings surprise as the difference between actual earnings per share and the last available consensus median forecast prior to the announcement of annual earnings, and identify firms that just meet or beat analysts' expectations as those reporting a surprise of 0¢ or 1¢. We focus on annual earnings because auditors' involvement with quarterly financial statements is limited. Following Burgstahler and Dichev (1997), we identify firms reporting small increases in earnings and small profits by scaling net income and the annual change in net income by beginning of year market value of common equity.[6] Untabulated results confirm a significant discontinuity at zero in the distribution of earnings surprises and earnings changes. However, in contrast to prior research, we find no evidence of a discontinuity in the distribution of earnings levels, either for all nonfinancial firms on Compustat in 2000 or for the sample firms.[7] Because zero earnings does not appear to be a relevant benchmark during the sample period, we focus on small earnings surprises (SURPRISE) and small increases in earnings (INCREASE), where these measures are indicator variables equal to 1 if the firm just meets or beats expectations or reports a small increase in earnings, respectively.

In addition to the auditor fee measures previously discussed, we include two variables as proxies for audit quality. Prior research suggests that Big 5 auditors are less likely to allow earnings management than non-Big 5 auditors (e.g., DeFond and Jiambalvo 1991, 1993; Becker et al. 1998; Francis et al. 1999), and that independence is decreasing in the length of auditor tenure (e.g., Beck et al. 1988b; Lys and Watts 1994). Thus, our analysis includes a Big 5 indicator variable (BIGFIVE) and a variable measuring the number of years the auditor has audited the financial statements of the company (AUDTEN).

We control for several factors that create incentives to meet or beat earnings benchmarks. Matsumoto (2002) reports that firms with high litigation risk, growth prospects, and institutional ownership are more likely to be concerned about missing earnings benchmarks. We measure litigation risk (LITRISK) as an indicator variable equal to 1 if the company operates in a high-risk industry, as identified by Francis et al. (1994). Growth is the market-to-book ratio (M/B), and institutional ownership (%INST) is the percentage of shares held by institutions (reported by Spectrum). In addition, Brown (2001) finds that loss firms are less likely to report positive earnings surprises, and thus our regression model includes an indicator variable (LOSS) equal to 1 if the firm reported a loss in 2000.

We control for operating cash flow scaled by average total assets (CFO) because firms with high cash flow may be more likely to beat earnings benchmarks and can more easily purchase outside consulting services. Acquisition and financing activity and poor performance are additional determinants of nonaudit fees (Firth 1997).[8] Thus, our model includes an indicator variable (FIN/ACQ) identifying firms that issued securities or made an acquisition during 2000 (reported by Securities Data Corporation). We include two measures of firm performance, the percentage compounded monthly return for 2000 adjusted for the CRSP value-weighted market index (ANNRET), and return on assets (ROA), defined as net income divided by average total assets. Finally, we control for firm size using the log of the market value of equity (LOGMVE).

---

[6] Consistent with Degeorge et al. (1999), the interval width is twice the interquartile range of the scaled earnings variable times the negative cube root of sample size. The resulting interval width is 0.02 for both the change in earnings and the level of earnings.

[7] We do not find significant results when we repeat the tests below for firms reporting small profits but, given the absence of a discontinuity in the distribution of earnings, this result is not surprising. Doubling the interval width to 0.04 produces a significant discontinuity; however, the interpretation of this finding is not clear. We also do not find significant results using the wider interval width to identify firms reporting small profits.

We estimate the following logit model, where BENCHMARK indicates the earnings benchmarks, SURPRISE or INCREASE, and FEEVAR indicates the alternative specifications of the fee variables, (1) FEERATIO, (2) RANKNON and RANKAUD, or (3) RANKTOT:

$$\text{Prob(BENCHMARK)} = F\begin{pmatrix} \beta_0 + \beta_1\text{FEEVAR} + \beta_2\text{BIGFIVE} + \beta_3\text{AUDTEN} + \beta_4\text{LITRISK} \\ + \beta_5\text{M/B} + \beta_6\text{LOGMVE} + \beta_7\text{\%INST} + \beta_8\text{LOSS} + \beta_9\text{CFO} \\ + \beta_{10}\text{FIN/ACQ} + \beta_{11}\text{ROA} + \beta_{12}\text{ANNRET} + u \end{pmatrix}. \quad (1)$$

Positive coefficients on FEERATIO or RANKNON are consistent with the hypothesis that the provision of nonaudit services to audit clients creates incentives for auditors to acquiesce to client pressure to manage earnings, while insignificant or negative coefficients suggest that the provision of nonaudit services has no effect on earnings management or reduces earnings management, respectively. A similar interpretation of the incentive effects of audit fees applies to the coefficient estimates on RANKAUD. Finally, RANKTOT tests for the potential effects of overall economic dependence in auditor-client contracting.

Our second approach to identifying earnings management uses discretionary accruals estimated with a cross-sectional modified Jones (1991) model (see DeFond and Jiambalvo 1994; Becker et al. 1998; Bartov et al. 2001). Discretionary accruals (DACC) are equal to:

$$\text{DACC} = \text{TA} - (\hat{\alpha} + \hat{\beta}_1[\Delta\text{REV} - \Delta\text{REC}] - \hat{\beta}_2\text{PPE}), \quad (2)$$

where all variables including the intercept are scaled by total assets at the beginning of the year, and TA is total accruals, defined as net income less cash from operations, ΔREV is the change in net revenues, ΔREC is the change in net receivables, and PPE is gross property, plant, and equipment. The estimates of α, β_1, and β_2 are obtained from estimating the following model at the industry level, where industry membership is identified using two-digit SIC codes:

$$\text{TA} = \alpha + \beta_1\Delta\text{REV} + \beta_2\text{PPE} + \varepsilon. \quad (3)$$

Using the absolute value of discretionary accruals (ABSDACC) to measure the combined effect of income-increasing and income-decreasing earnings management decisions (Warfield et al. 1995; Becker et al. 1998; Reynolds and Francis 2000), we estimate the following model:

---

[8] In the next section, we discuss the estimation of a model of nonaudit fees similar to that of Firth (1997). The results of this analysis are reported in the Appendix.

$$\begin{aligned} \text{ABSDACC} = \alpha &+ \beta_1\text{FEEVAR} + \beta_2\text{BIGFIVE} + \beta_3\text{AUDTEN} + \beta_4\text{CFO} \\ &+ \beta_5\text{ABSCFO} + \beta_6\text{ACC} + \beta_7\text{ABSACC} + \beta_8\text{LEVERAGE} \\ &+ \beta_9\text{LITRISK} + \beta_{10}\text{M/B} + \beta_{11}\text{LOGMVE} + \beta_{12}\%\text{INST} \\ &+ \beta_{13}\text{LOSS} + \beta_{14}\text{FIN/ACQ} + \beta_{15}\text{ANNRET} + \varepsilon. \end{aligned} \tag{4}$$

The predictions for the coefficient estimates on the fee variables are as discussed previously for Equation (1). We also present separate results for firms with income-increasing discretionary accruals (DACC^+) and income-decreasing discretionary accruals (DACC^-).

Based on prior research showing that discretionary accruals models do not completely extract nondiscretionary accruals that are correlated with firm performance (e.g., Dechow et al. 1995; McNichols 2000), we include four measures of firm performance, all of which are deflated by average total assets: cash from operations (CFO), the absolute value of cash from operations (ABSCFO), total accruals (ACC), and the absolute value of total accruals (ABSACC). We also control for leverage (LEVERAGE), measured as the ratio of total liabilities to total assets, because prior research finds leverage is associated with discretionary accruals (DeFond and Jiambalvo 1994; DeAngelo et al. 1994; Becker et al. 1998). The remaining control variables in Equation (4) are as previously defined.

Descriptive Statistics

Table 4, Panel A reports descriptive statistics for the sample partitioned at the median of FEERATIO, which, as shown in Panel A of Table 2, is 0.51. The findings reveal that firms with a high FEERATIO are also of greater financial importance to their auditors, as measured by RANKNON, RANKAUD, and RANKTOT. In addition, firms with a high FEERATIO are more likely to have a Big 5 auditor, consistent with the findings reported in Table 3. All three earnings management indicators, SURPRISE, INCREASE, and ABSDACC, are significantly higher for firms in the high FEERATIO partition.

The remaining variables in Table 4 indicate several significant differences in the economic characteristics of firms in the two FEERATIO subsamples. High FEERATIO firms have higher cash flows and profitability, lower total accruals and stock performance, higher litigation risk, growth, and institutional shareholdings, and are more likely to engage in financing and acquisition activities. Finally, high FEERATIO firms are significantly larger than other firms. The empirical tests that follow control for these factors.

Table 4, Panel B reports correlations for the primary regression variables and firm size. We do not tabulate correlations for the other regression variables as they are generally small, i.e., less than ±0.30. The results indicate that FEERATIO is highly positively correlated with RANKNON and RANKTOT, but less so with RANKAUD. Consistent with prior research (e.g., Simunic 1984; Palmrose 1986; Craswell et al. 1995), we find a positive correlation between RANKNON and RANKAUD. All four fee variables are positively correlated with SURPRISE, INCREASE, and LOGMVE. In contrast, ABSDACC is positively correlated with FEERATIO, but negatively correlated with the other fee variables.

IV. EMPIRICAL RESULTS

Auditor Fees and Earnings Management

Benchmark Tests

Our first test examines the association between earnings management, as indicated by just meeting or beating earnings benchmarks, and auditor fees. We report results for three

TABLE 4
Sample Descriptive Statistics

Panel A: Comparison of Distributions for Regression Variables Partitioned at the Median Value of FEERATIO

	Below Median FEERATIO				*Above Median FEERATIO*				*p-value*	
Variable	*n*	*Mean*	*Median*	*Standard Deviation*	*n*	*Mean*	*Median*	*Standard Deviation*	*t-test*	*Wilcoxon*
RANKNON	1,537	30.42	26.00	21.70	1,537	68.61	71.00	20.08	<0.01	<0.01
RANKAUD	1,537	43.19	43.00	26.70	1,537	55.85	58.00	28.50	<0.01	<0.01
RANKTOT	1,537	34.31	30.00	24.33	1,537	64.73	68.00	23.45	<0.01	<0.01
BIGFIVE	1,537	0.84	1.00	0.37	1,537	0.97	1.00	0.17	<0.01	<0.01
AUDTEN	1,528	8.13	6.00	5.90	1,528	8.36	6.00	6.31	0.31	0.94
SURPRISE	936	0.23	0.00	0.42	1,316	0.31	0.00	0.46	<0.01	<0.01
INCREASE	1,537	0.14	0.00	0.34	1,537	0.20	0.00	0.40	<0.01	<0.01
ABSDACC	1,430	0.12	0.07	0.20	1,383	0.17	0.07	0.33	<0.01	0.05
CFO	1,530	−0.05	0.04	0.31	1,529	0.01	0.06	0.23	<0.01	<0.01
ABSCFO	1,530	0.18	0.10	0.26	1,529	0.15	0.10	0.17	<0.01	0.54
ACC	1,530	−0.08	−0.05	0.21	1,529	−0.10	−0.06	0.30	0.14	0.04
ABSACC	1,530	0.01	0.01	0.16	1,529	0.00	0.00	0.01	<0.01	<0.01
LITRISK	1,537	0.36	0.00	0.48	1,537	0.43	0.00	0.50	<0.01	<0.01
M/B	1,424	3.16	1.45	6.27	1,475	3.69	2.13	6.45	0.02	<0.01
%INST	1,414	0.30	0.24	0.26	1,475	0.45	0.45	0.26	<0.01	<0.01
LOSS	1,537	0.47	0.00	0.50	1,537	0.42	0.00	0.49	<0.01	<0.01
FIN/ACQ	1,473	0.29	0.00	0.46	1,427	0.51	1.00	0.50	<0.01	<0.01
LEVERAGE	1,536	0.54	0.49	0.48	1,533	0.47	0.47	0.28	0.87	0.16
ROA	1,534	−0.14	0.01	0.42	1,533	−0.10	0.02	0.42	0.01	<0.01
ANNRET	1,342	0.42	0.29	0.47	1,353	0.32	0.22	0.31	<0.01	<0.01
MVE	1,530	822.30	65.40	8,753.28	1,518	4,923.87	483.92	21,951.39	<0.01	<0.01
ASSET	1,537	675.09	95.28	3,059.70	1,537	3,997.40	432.68	18,715.02	<0.01	<0.01

(*Continued on next page*)

TABLE 4 (Continued)

Panel B: Pearson (Spearman) Correlation Coefficients above (below) the Diagonal

	FEERATIO	*RANKNON*	*RANKAUD*	*RANKTOT*	*SURPRISE*	*INCREASE*	*ABSDACC*	*LOGMVE*
FEERATIO		0.77	0.24	0.60	0.10	0.09	0.09	0.50
RANKNON	0.77		0.73	0.95	0.11	0.12	−0.01*	0.61
RANKAUD	0.23	0.72		0.88	0.07	0.12	−0.12	0.55
RANKTOT	0.60	0.95	0.88		0.10	0.13	−0.03*	0.62
SURPRISE	0.10	0.11	0.07	0.10		0.15	−0.03*	0.18
INCREASE	0.08	0.12	0.12	0.13	0.15		−0.12	0.31
ABSDACC	0.03	−0.09	−0.21	−0.14	−0.03*	−0.16		−0.06
LOGMVE	0.50	0.60	0.53	0.61	0.17	0.29	−0.16	

(Continued on next page)

TABLE 4 (Continued)

---

* Denotes a significance level > 0.10. All other correlations are significant at the 0.10 level or below.

FEERATIO = ratio of financial information system and other fees to total fees;
RANKNON = percentile rank of nonaudit fees, by auditor;
RANKAUD = percentile rank of audit fees, by auditor;
RANKTOT = percentile rank of total fees, by auditor;
BIGFIVE = 1 if the firm's auditor is a Big 5 firm, and 0 otherwise;
AUDTEN = number of years that the auditor has audited the firm's financial statements;
SURPRISE = 1 if the firm just meets or beats the consensus analyst forecast (i.e., forecast error of 0¢ or 1¢) for fiscal 2000, and 0 otherwise;
INCREASE = 1 if the firm reports a small increase in earnings relative to the prior year, and 0 otherwise;
ABSACC = absolute value of total accruals, equal to net income minus cash from operations, deflated by average total assets;
CFO = cash from operations, deflated by average total assets;
ABSCFO = absolute value of cash from operations deflated by average total assets;
ACC = total accruals, equal to net income minus cash from operations, deflated by average total assets;
LITRISK = 1 if the firm is in a high litigation risk industry identified by Francis et al. (1994) (SICs 2833–2836, 3570–3577, 7370–7374, 3600–3674, 5200–5961), and 0 otherwise;
M/B = market-to-book ratio;
%INST = percent of shares held by institutions (as reported by Spectrum);
LOSS = 1 if the firm reported a net loss in fiscal 2000, and 0 otherwise;
FIN/ACQ = 1 if the firm issued securities or acquired another company in fiscal 2000, and 0 otherwise;
LEVERAGE = ratio of total liabilities to total assets;
ROA = net income divided by average total assets;
ANNRET = percentage compounded monthly stock return for calendar 2000 minus the CRSP value-weighted market index;
MVE = market value of equity (in millions);
LOGMVE = natural log of MVE; and
ASSET = book value of total assets (in millions).

specifications of auditor fees: (1) FEERATIO, (2) RANKNON and RANKAUD, and (3) RANKTOT. The findings, reported in Table 5, are mixed.[9]

The associations between SURPRISE and the nonaudit fee measures FEERATIO and RANKNON are positive and significant at the 0.01 level. These results suggest that firms purchasing more nonaudit services are more likely to just meet analysts' expectations, consistent with the earnings management concerns expressed by the SEC (2000) and POB (2000). In contrast, the coefficient estimate on RANKAUD is negative and significant, suggesting that high audit fees are associated with less earnings management.[10] Finally, the coefficient on RANKTOT is insignificant (p-value = 0.62). This finding, in conjunction with the results for RANKNON and RANKAUD, indicates that the composition of fees paid to the auditor is relevant, but the total amount of fees is not. In all three specifications of the SURPRISE regression, neither of the other two auditor-related variables, BIGFIVE and AUDTEN, is significant at the 0.10 level. Other results indicate that firms reporting small positive earnings surprises have greater litigation risk and institutional ownership, are larger, and are less likely to report a loss.

Untabulated findings indicate that the results are robust to an earnings surprise indicator variable for firms (1) reporting a small positive surprise in all four quarters of fiscal 2000, or (2) reporting a small positive surprise in at least three quarters of fiscal 2000. The coefficient estimates on the fee variables are insignificant at traditional levels for firms reporting a small positive surprise in at least two quarters of fiscal 2000. However, firms reporting small positive surprises in only two quarters likely have less incentive to manage earnings to meet or beat expectations, and may not be doing so for audited annual earnings.[11] Consistent with the latter argument, we find that FEERATIO and RANKNON are positive and significant and RANKAUD is negative and significant at the 0.10 level or less for firms reporting a small positive surprise in at least two quarters of fiscal 2000, provided at least one of the quarters is the fourth quarter.[12]

The basic premise of the benchmark analysis is that earnings management occurs in firms that just meet or beat the benchmark but not in firms that just miss the benchmark. Based on this premise, we estimate the SURPRISE model with the dependent variable defined as an indicator variable equal to 1 for firms with forecast errors of −1¢ or −2¢. The untabulated results indicate that FEERATIO and RANKNON are negative and significant (p-value < 0.01); RANKAUD is insignificant at conventional levels.[13] These results

---

[9] All results are presented after the deletion of statistical outliers. However, inferences are qualitatively similar if outliers are retained.

[10] We also estimate the model separately for RANKNON and RANKAUD. The untabulated results indicate that RANKNON is positively associated with small positive earnings surprises (p-value = 0.09). The coefficient on RANKAUD is negative but not statistically significant at conventional levels (p-value = 0.19). Because RANKNON and RANKAUD are positively correlated (see Table 4), but have coefficient estimates of the opposite sign, omitting one of the variables biases the coefficient of the included variable toward zero.

[11] In our sample, 37 percent of the firms that report two or three quarters of small positive earnings surprises do not report a small positive surprise in the fourth quarter.

[12] We also examine the robustness of the results using the mean consensus forecast to calculate the earnings surprise. The untabulated results indicate that FEERATIO and RANKNON are less significant (p-values of 0.09 and 0.11, respectively), perhaps due to outliers causing error in the mean forecast as a measure of market expectations. RANKAUD remains significant at the 0.01 level.

[13] Untabulated findings also indicate that our results are robust to defining the earnings surprise indicator variable to be equal to 1 if the firm reports a small positive earnings surprise and 0 if the firm reports a small negative earnings surprise. Francis and Ke (2002) argue that this specification of the dependent variable is more consistent with Degeorge et al. (1999). Using data from all four quarters of fiscal 2000, Francis and Ke (2002) do not find a significant positive coefficient estimate on FEERATIO. Because our findings are not sensitive to the alternative specification of the dependent variable, these results suggest, consistent with the evidence documented above, that including un-audited quarters in the analysis reduces the ability of the test to identify an association between nonaudit fees and earnings management to meet analysts' earnings forecasts.

TABLE 5
Logit Regressions Modeling the Probability of Meeting Earnings Benchmarks

$$Prob(BENCHMARK) = F\begin{pmatrix} \beta_0 + \beta_1 FEEVAR + \beta_2 BIGFIVE + \beta_3 AUDTEN + \beta_4 LITRISK + \beta_5 M/B + \beta_6 LOGMVE \\ + \beta_7 \%INST + \beta_8 LOSS + \beta_9 CFO + \beta_{10} FIN/ACQ + \beta_{11} ROA + \beta_{12} ANNRET + u \end{pmatrix}$$

	BENCHMARK											
	SURPRISE						*INCREASE*					
Variable	*Coeff. Est.*	*p-value*	*Coeff. Est.*	*p-value*	*Coeff. Est.*	*p-value*	*Coeff. Est.*	*p-value*	*Coeff. Est.*	*p-value*	*Coeff. Est.*	*p-value*
Intercept	−1.37	<0.01	−1.28	<0.01	−1.32	<0.01	−1.04	<0.01	−0.91	<0.01	−0.96	<0.01
FEERATIO	0.42	<0.01					−0.07	0.72				
RANKNON			0.01	<0.01					−0.01	0.83		
RANKAUD			−0.01	<0.01					−0.01	0.02		
RANKTOT					0.01	0.62					−0.01	0.02
BIGFIVE	−0.01	0.94	0.05	0.77	0.05	0.75	−0.20	0.24	−0.29	0.12	−0.28	0.13
AUDTEN	−0.01	0.65	−0.01	0.88	−0.01	0.50	−0.01	0.35	−0.01	0.75	−0.01	0.58
LITRISK	0.26	<0.01	0.24	<0.01	0.26	<0.01	0.19	0.04	0.15	0.10	0.17	0.07
M/B	−0.01	0.99	−0.01	0.88	−0.01	0.91	−0.01	0.61	−0.01	0.43	−0.01	0.44
LOGMVE	0.07	<0.01	0.07	<0.01	0.08	<0.01	0.16	<0.01	0.21	<0.01	0.20	<0.01
%INST	0.40	<0.01	0.41	<0.01	0.39	<0.01	0.03	0.89	0.05	0.78	0.05	0.80
LOSS	−0.34	<0.01	−0.34	<0.01	−0.33	<0.01	−1.62	<0.01	−1.79	<0.01	−1.75	<0.01
CFO	0.06	0.83	0.04	0.90	0.04	0.90	−0.02	0.97	−0.10	0.96	0.01	0.98
FIN/ACQ	−0.02	0.66	−0.01	0.87	−0.02	0.79	−0.14	0.08	−0.11	0.22	−0.10	0.21
ROA	0.34	0.18	0.32	0.20	0.31	0.22	0.40	0.52	0.22	0.74	0.33	0.61
ANNRET	0.15	0.32	0.13	0.41	0.17	0.27	−2.09	<0.01	−2.42	0.01	−2.35	<0.01
Pseudo R^2	0.07		0.07		0.06		0.19		0.20		0.20	
Nobs.	2,012		2,014		2,014		2,092		2,090		2,090	

The variables are defined in Table 4, where FEEVAR indicates the alternative specifications of the fee variables, (1) FEERATIO, (2) RANKNON and RANKAUD, or (3) RANKTOT.

provide additional evidence consistent with an association between nonaudit fees and earnings management to meet or slightly exceed analysts' expectations.

Contrary to the findings for the SURPRISE model, we find no evidence that firms paying high nonaudit fees are more likely to report small earnings increases. Specifically, both FEERATIO and RANKNON are insignificant in the INCREASE regressions. Our results are thus consistent with Gore et al. (2001) who find a positive association between the purchase of nonaudit services and meeting analysts' expectations for U.K. firms with a Big 5 auditor, but no association between nonaudit services and reporting small earnings increases.[14] Francis and Ke (2002) also find no relation between small earnings increases and the ratio of nonaudit fees to total fees using U.S. data. In contrast to the results for nonaudit fees, we find that RANKAUD is negatively associated with small earnings increases, as is RANKTOT. The results for the control variables in the INCREASE regressions are generally consistent with the SURPRISE regressions with two exceptions; sample firms reporting a small increase in earnings have poor market performance and do not attract institutional owners.

One possible explanation for the different findings in the SURPRISE and INCREASE regressions is that managers have greater incentives to manage earnings to meet market expectations and/or auditors have less incentives to prevent this behavior. Consistent with this explanation, Nelson et al. (2002) report that 17 percent of earnings management attempts are to meet analyst expectations, and that auditors frequently waive these attempts (72 percent of the time). In contrast, 3 percent of earnings management attempts are to show an increase in earnings, and auditors are less likely to waive these attempts (50 percent of the time).

Discretionary Accruals Tests

Our second test examines the relation between discretionary accruals as an indicator of earnings management and auditor fees. Table 6 reports summary statistics for the absolute value of discretionary accruals (Panel A) and for separate partitions of income-increasing and income-decreasing discretionary accruals (Panel B).[15] The results in Panel A for FEERATIO and RANKNON indicate that purchases of nonaudit services are positively associated with the magnitude of discretionary accruals, while the results for RANKAUD indicate the opposite is true for purchases of audit services. The coefficient estimate on RANKTOT is insignificant (p-value = 0.16). In related work, Chung and Kallapur (2001) deflate fees by the audit firm's total U.S. revenue, and find no association between absolute discretionary accruals and the scaled measures of either total fees or nonaudit fees. The difference in findings for nonaudit fees is consistent with the scaled fee data measuring the financial importance of the client with error.

Consistent with the inferences for the absolute discretionary accruals regressions, the separate estimations for income-increasing and income-decreasing discretionary accruals reported in Panel B of Table 6 reveal that FEERATIO and RANKNON are positive and significant in the $DACC^+$ regressions and negative and significant in the $DACC^-$ regressions. RANKAUD is also significant in both regressions with a sign opposite that of the nonaudit fee variables. RANKTOT is positive and significant (p-value = 0.06) in the $DACC^+$ model, but is insignificant (p-value = 0.18) in the $DACC^-$ model.

---

[14] Although our sample includes clients of non-Big 5 auditors, 96 percent of the firms included in the Table 5 regressions are audited by a Big 5 auditor.

[15] All results are robust to t-statistics calculated using White's (1980) consistent standard error estimates to correct for heteroskedasticity.

TABLE 6
Summary Statistics from Discretionary Accruals Regressions

Panel A: Dependent Variable Is Absolute Value of Discretionary Accruals

$$ABSDACC = \alpha + \beta_1 FEEVAR + \beta_2 BIGFIVE + \beta_3 AUDTEN + \beta_4 CFO + \beta_5 ABSCFO + \beta_6 ACC + \beta_7 ABSACC + \beta_8 LEVERAGE + \beta_9 LITRISK + \beta_{10} M/B + \beta_{11} LOGMVE + \beta_{12} \%INST + \beta_{13} LOSS + \beta_{14} FIN/ACQ + \beta_{15} ANNRET + \varepsilon$$

Variable	*Coefficient Estimate*	*p-value*	*Coefficient Estimate*	*p-value*	*Coefficient Estimate*	*p-value*
Intercept	0.06	<0.01	0.06	<0.01	0.06	<0.01
FEERATIO	0.07	<0.01				
RANKNON			0.01	<0.01		
RANKAUD			−0.01	<0.01		
RANKTOT					0.01	0.16
BIGFIVE	−0.01	0.65	−0.01	0.97	0.01	0.68
AUDTEN	−0.01	0.03	−0.01	0.07	−0.01	0.01
CFO	−0.19	<0.01	−0.20	<0.01	−0.20	<0.01
ABSCFO	−0.26	<0.01	−0.28	<0.01	−0.27	<0.01
ACC	0.16	<0.01	0.16	<0.01	0.16	<0.01
ABSACC	1.17	<0.01	1.18	<0.01	1.17	<0.01
LEVERAGE	−0.16	<0.01	−0.14	<0.01	−0.16	<0.01
LITRISK	0.01	0.39	0.01	0.45	0.01	0.28
M/B	0.01	0.39	0.01	0.83	0.01	0.44
LOGMVE	0.01	<0.01	0.02	<0.01	0.01	<0.01
%INST	−0.05	<0.01	−0.05	<0.01	−0.05	<0.01
LOSS	−0.02	0.06	−0.02	0.04	−0.02	0.08
FIN/ACQ	0.01	0.98	0.01	0.64	0.01	0.87
ANNRET	0.03	0.23	0.02	0.20	0.02	0.22
Adjusted R^2	0.46		0.47		0.46	
n	2,472		2,472		2,472	

(Continued on next page)

TABLE 6 (Continued)

Panel B: Dependent Variable Is Signed Discretionary Accruals

$$DACC = \alpha + \beta_1 FEEVAR + \beta_2 BIGFIVE + \beta_3 AUDTEN + \beta_4 CFO + \beta_5 ABSCFO + \beta_6 ACC + \beta_7 ABSACC + \beta_8 LEVERAGE + \beta_9 LITRISK + \beta_{10} M/B + \beta_{11} LOGMVE + \beta_{12} \%INST + \beta_{13} LOSS + \beta_{14} FIN/ACQ + \beta_{15} ANNRET + \varepsilon$$

	$DACC^+$						$DACC^-$					
Variable	*Coeff. Est.*	*p-value*	*Coeff. Est.*	*p-value*	*Coeff. Est.*	*p-value*	*Coeff. Est.*	*p-value*	*Coeff. Est.*	*p-value*	*Coeff. Est.*	*p-value*
Intercept	0.04	<0.01	0.05	<0.01	0.05	<0.01	0.02	0.56	0.01	0.78	0.01	0.87
FEERATIO	0.05	<0.01					−0.10	<0.01				
RANKNON			0.01	<0.01					−0.01	<0.01		
RANKAUD			−0.01	<0.01					0.01	<0.01		
RANKTOT					0.01	0.06					−0.01	0.18
BIGFIVE	0.01	0.68	0.01	0.21	0.01	0.14	−0.01	0.91	−0.01	0.63	−0.02	0.45
AUDTEN	−0.01	0.27	−0.01	0.29	−0.01	0.14	0.01	0.01	0.01	0.02	0.01	<0.01
CFO	0.01	0.88	0.01	0.95	0.01	0.93	0.17	<0.01	0.18	<0.01	0.17	<0.01
ABSCFO	0.02	0.43	0.02	0.51	0.02	0.45	0.28	<0.01	0.30	<0.01	0.29	<0.01
ACC	0.58	<0.01	0.58	<0.01	0.58	<0.01	−0.35	0.75	−0.33	0.77	−0.42	0.71
ABSACC	0.51	<0.01	0.51	<0.01	0.51	<0.01	−1.50	0.18	−1.50	0.19	−1.58	0.17
LEVERAGE	−0.03	0.01	−0.03	0.05	−0.04	<0.01	0.20	<0.01	0.18	<0.01	0.22	<0.01
LITRISK	0.01	0.03	0.01	0.04	0.01	0.01	−0.01	0.66	−0.01	0.73	−0.01	0.63
M/B	0.01	0.76	0.01	0.93	0.01	0.69	0.01	0.97	0.01	0.64	0.01	0.99
LOGMVE	0.01	0.86	0.01	0.63	0.01	0.77	−0.03	<0.01	−0.03	<0.01	−0.03	<0.01
%INST	−0.02	0.16	−0.02	0.21	−0.02	0.14	0.08	0.01	0.08	0.02	0.08	0.02
LOSS	0.01	0.51	0.01	0.51	0.01	0.48	0.01	0.51	0.01	0.68	0.01	0.87
FIN/ACQ	0.01	0.26	0.01	0.20	0.01	0.22	0.01	0.56	0.01	0.86	0.01	0.59
RETANN	0.05	<0.01	0.05	<0.01	0.05	<0.01	−0.01	0.88	−0.58	0.79	−0.34	0.88
Adjusted R^2	0.40		0.40		0.39		0.52		0.52		0.51	
n	1,355		1,355		1,355		1,116		1,116		1,116	

The variables are defined in Table 4, where FEEVAR indicates the alternative specifications of the fee variables, (1) FEERATIO, (2) RANKNON and RANKAUD, or (3) RANKTOT.

We repeat the Table 6 analyses using discretionary measures based on working capital accruals, defined as total accruals less depreciation expense, and operating accruals, defined as income before special items less cash from operations. All inferences are unchanged. We also examine a measure of discretionary accruals based on Kothari et al. (2002) who adjust discretionary accruals of treatment firms by subtracting the discretionary accrual of an industry and performance-matched control firm. Because 96 percent of our sample purchased nonaudit services (see Table 2), leaving few control firms, we modify the Kothari et al. (2002) procedure by partitioning firms in each one-digit SIC code into deciles based on lagged return on assets (ROA). Performance-matched discretionary accruals are equal to the sample firm's discretionary accruals less the discretionary accruals of the firm in the same industry/ROA decile with the lowest FEERATIO. This procedure identifies industry- and performance-matched control firms purchasing few or no nonaudit services.[16] We repeat the analyses in Table 6 with performance-matched discretionary accruals for both total and working capital accruals, and find that FEERATIO and RANKNON are consistently positive and significant at the 0.10 level or less; RANKAUD is negative and significant except in the DACC$^+$ estimations.[17]

Additional Sensitivity Analysis

To examine the effects of variation in the composition of fees across auditors (see Table 3) on our results, we repeat the analyses in Tables 5 and 6 separately for each of the Big 5 auditors and for the non-Big 5 auditors as a group. Table 7 summarizes the findings for the three sets of auditor fee variables: (1) FEERATIO, (2) RANKNON and RANKAUD, and (3) RANKTOT. With few exceptions, the signs of the coefficient estimates are consistent with the pooled results. Moreover, many of the insignificant coefficients have p-values between 0.11 and 0.15. The results for three firms, EY, KPMG, and PWC, are generally more significant than the results for the other firms.[18] One possible explanation for this finding is that these firms have higher percentages of nonaudit revenue, as shown in Table 3, and thus have greater incentives to acquiesce to client pressure to manage earnings. Alternatively, these results may reflect differences in the types of nonaudit services provided and their associated incentive effects, or cross-firm differences in the interpretation of fees to include in each category.

To control for potential nonlinearities in the relation between firm size and the incentive effects of auditor fees, we partition the sample into deciles based on LOGMVE and estimate

---

[16] The use of one-digit SIC codes represents a trade-off between defining industry groupings narrowly enough that the matching procedure captures industry-specific effects and having enough observations in each partition to be able to identify a control firm that differs from the treatment firms. Mean (median) FEERATIO for the treatment firms is 0.52 (0.53), compared to 0.03 (0.00) for the control firms; both of these differences are significant at less than the 0.01 level. Mean (median) ROA is −0.08 (0.02) for the treatment firms, compared to −0.04 (0.02) for the control firms. Although the mean difference is significant at the 0.08 level, the median difference is insignificant (p-value = 0.74).

[17] We repeat the analyses using a control group matched at the two-digit SIC level. This approach substantially increases the treatment effect found in the control group, as mean (median) FEERATIO is 0.21 (0.19). This approach also weakens the performance match, as the mean (median) difference in ROA between the treatment and control firms is significant at the 0.01 level. As a consequence, the relation between performance-matched discretionary accruals and the auditor fee variables is less significant, particularly for income-increasing discretionary accruals where none of the fee variables is significant at conventional levels. The latter findings are consistent with Ashbaugh et al. (2002) who calculate performance-matched discretionary accruals using the median discretionary accrual firm in each industry/ROA decile as the control for that decile. This procedure results in a control group with approximately the same level of nonaudit fees as the treatment group; i.e., for our sample, mean (median) FEERATIO is 0.51 (0.53) for the treatment firms compared to 0.48 (0.49) for the control firms.

[18] Untabulated findings indicate that FEERATIO, RANKNON, and RANKAUD are significant at the 0.10 level or less for the Big 5 sample as a group; RANKTOT is insignificant.

TABLE 7
Summary of Results from Auditor-Specific Regressions

		AA		*DT*		*EY*		*KPMG*		*PWC*		*All Other*	
Variable		*Coeff. Est.*	*p-value*	*Coeff. Est.*	*p-value*	*Coeff. Est.*	*p-value*	*Coeff. Est.*	*p-value*	*Coeff. Est.*	*p-value*	*Coeff. Est.*	*p-value*
Panel A: SURPRISE Regressions													
Regr. 1:	FEERATIO	0.17	0.60	−0.04	0.91	**1.25**	**<0.01**	0.09	0.83	0.51	0.12	0.59	0.53
Regr. 2:	RANKNON	0.01	0.55	0.01	0.93	**0.01**	**<0.01**	−0.01	0.95	0.01	0.12	0.01	0.19
	RANKAUD	−0.01	0.54	**−0.01**	**0.07**	**−0.01**	**0.09**	−0.01	0.22	−0.01	0.39	**−0.03**	**0.01**
Regr 3:	RANKTOT	0.01	0.92	−0.01	0.13	**0.01**	**0.03**	−0.01	0.19	0.01	0.35	−0.01	0.15
	n	423		260		475		285		488		83	
Panel B: ABSDACC Regressions													
Regr. 1:	FEERATIO	0.01	0.73	0.01	0.69	**0.10**	**<0.01**	**0.09**	**0.08**	**0.10**	**<0.01**	0.05	0.31
Regr. 2:	RANKNON	0.01	0.75	0.01	0.63	**0.01**	**<0.01**	0.01	0.12	**0.01**	**<0.01**	−0.01	0.79
	RANKAUD	−0.01	0.17	−0.01	0.90	**−0.01**	**<0.01**	−0.01	0.18	**−0.01**	**0.02**	**−0.01**	**0.07**
Regr 3:	RANKTOT	−0.01	0.54	0.01	0.73	0.01	0.12	0.01	0.35	0.01	0.18	**−0.01**	**0.08**
	n	488		307		582		336		565		194	
Panel C: $DACC^+$ Regressions													
Regr. 1:	FEERATIO	0.01	0.73	0.02	0.32	**0.05**	**0.04**	**0.14**	**<0.01**	0.04	0.11	0.03	0.54
Regr. 2:	RANKNON	0.01	0.95	0.01	0.13	**0.01**	**0.05**	**0.01**	**<0.01**	**0.01**	**0.06**	−0.01	0.86
	RANKAUD	−0.01	0.78	−0.01	0.66	**−0.01**	**0.01**	−0.01	0.13	0.01	0.26	**−0.01**	**0.03**
Regr 3:	RANKTOT	−0.01	0.77	0.01	0.23	−0.01	0.90	**0.01**	**0.02**	**0.01**	**<0.01**	**−0.01**	**0.01**
	n	247		183		318		188		297		122	

(*Continued on next page*)

TABLE 7 (Continued)

Variable		AA		DT		EY		KPMG		PWC		All Other	
		Coeff. Est.	*p-value*	*Coeff. Est.*	*p-value*	*Coeff. Est.*	*p-value*	*Coeff. Est.*	*p-value*	*Coeff. Est.*	*p-value*	*Coeff. Est.*	*p-value*
Panel D: $DACC^-$ Regressions													
Regr. 1:	FEERATIO	−0.04	0.39	−0.04	0.45	**−0.15**	**0.03**	−0.01	0.90	**−0.17**	**0.02**	−0.11	0.21
Regr. 2:	RANKNON	−0.01	0.43	−0.01	0.71	**−0.01**	**<0.01**	0.01	0.76	**−0.01**	**0.01**	−0.01	0.45
	RANKAUD	0.01	0.28	0.01	0.95	**0.01**	**0.03**	0.01	0.67	**0.01**	**0.02**	0.01	0.32
Regr 3:	RANKTOT	−0.01	0.84	−0.01	0.79	−0.01	0.12	0.01	0.67	−0.01	0.32	0.01	0.91
	n	241		124		264		148		268		72	

This table summarizes the results from auditor-specific regressions of alternative measures of earnings management (SURPRISE, ABSDACC, $DACC^+$, or $DACC^-$) on FEERATIO (Regr. 1), RANKNON and RANKAUD (Regr. 2) and RANKTOT (Regr. 3), and control variables.

The variables are defined in Table 4.

The regression equations are defined in Tables 5 and 6.

For parsimony, summary statistics for the control variables are not tabulated. The auditors are Arthur Andersen (AA), Deloitte & Touche (DT), Ernst & Young (EY), KPMG (KPMG), PricewaterhouseCoopers (PWC), and all other auditors are combined in the category "All Other."

Coefficient estimates significant at the 0.10 level or below are highlighted in boldface type.

the earnings surprise and discretionary accruals models for each decile.[19] In the discretionary accruals regressions, we find no evidence of a size effect for FEERATIO and RANKNON. Specifically, the signs of the coefficient estimates are generally consistent with the pooled results and there is no evidence of a pattern in the estimates across the deciles. In contrast, there is evidence of a size effect for RANKAUD in the absolute and positive discretionary accruals estimations. In both cases, the coefficient estimates on RANKAUD are less negative for larger firms. The analysis also indicates a size effect in the SURPRISE estimations, as FEERATIO, RANKNON, and RANKAUD are not significant at the 0.10 level in the upper half of the sample.

We also disaggregate nonaudit fees into IS fees and other fees. Because of concerns that the provision of IS services poses a particularly strong threat to independence, the preliminary independence rule prohibited auditors from providing these services to their audit clients. The final Rule relaxed the ban, but requires separate disclosure of IS fees. We repeat the analyses in Tables 5 and 6 using IS fees and other fees, measured either by deflating by total fees or as percentile ranks. Only the measures of other fees are consistently significant at the 0.10 level or less, possibly because relatively few firms report IS fees. Finally, to control for systematic industry differences, we include fixed-effects for the industries identified in Table 1. We also repeat the analyses using the intersection of 1,943 firms that are common to both the benchmark and discretionary accruals tests. The results from both of these specifications are consistent with our primary findings.

Market Reaction to the Disclosure of Auditor Fees

We examine the market reaction to the disclosure of auditor fees using an event study. To identify the earliest date at which the fee information became public, we search EDGAR for firms that filed a preliminary proxy disclosing auditor fees.[20] If the firm filed a preliminary proxy disclosing auditor fees, then we use the filing date of the preliminary proxy as the event date (day 0); if not, we use the filing date of the definitive proxy. To mitigate the effects of confounding events, we eliminate firms that filed a 10-K, 10-Q, or 8-K within two days of the proxy.[21]

Using data from the 2001 CRSP tape, we calculate abnormal returns (ARET) as event date market model prediction errors, estimated from firm-level regressions of raw returns on NYSE size-decile portfolio returns over the 100-day period ending 30 days prior to the event. Because stock prices will react to any new information contained in the fee disclosures, we require a proxy for the unexpected component of auditor fees. If the fees were entirely or mostly unexpected, as suggested by anecdotal evidence (e.g., Weil and Tannenbaum 2001), then the fee variables discussed in Section III are reasonable proxies. We also estimate the unexpected component of nonaudit fees, scaled by total fees, using a model based on Parkash and Venable (1993) and Firth (1997) that assumes nonaudit fees vary with agency costs, among other factors. The Appendix describes this model. The residual from this model (FEERESIDUAL) provides an alternative measure of unexpected nonaudit fees.

---

[19] These estimations also include LOGMVE as a regressor to control for cross-sectional variation in size within deciles. The coefficient estimate on LOGMVE is generally not significant at the 0.10 level except in the largest size decile.

[20] Companies are required to file a preliminary proxy for review by the SEC if the annual meeting will involve voting on nonroutine matters. Thus, preliminary proxies are more likely to contain value-relevant information beyond the auditor fee disclosures. However, our results are robust to eliminating the approximately 10 percent (314 firms) of the sample filing preliminary proxy statements.

[21] Approximately 20 percent (614 firms) of the sample was deleted for this reason. The results are less significant when these observations are included.

We estimate the following model, where FEEVAR indicates the fee variables, (1) FEERATIO, (2) RANKNON and RANKAUD, (3) RANKTOT, or (4) FEERESIDUAL, and firm size (LOGMVE) controls for residual firm size effects:

$$ARET = \beta_0 + \beta_1 FEEVAR + \beta_2 LOGMVE + \varepsilon. \quad (5)$$

Table 8 reports summary statistics from the estimation of Equation (5).[22] For each fee variable, we tabulate results both including and excluding the firm size control. The coefficient estimates on all three measures of unexpected nonaudit fees, FEERATIO, RANKNON, and FEERESIDUAL, are negative and significant at the 0.10 level or below, as is the coefficient estimate on unexpected total fees, RANKTOT. Moreover, including firm size in the estimation improves the significance level of the coefficient estimates on FEERATIO, RANKNON, and RANKTOT to 0.01 or below, consistent with firm size acting as an important control in measuring unexpected nonaudit fees. Because FEERESIDUAL is orthogonal to firm size, it is not surprising that LOGMVE is insignificant (p-value = 0.20) in that regression. The coefficient estimates on RANKAUD are insignificant (p-value = 0.81) whether or not size is included.

We examine the robustness of these results by expanding the event window to two days, i.e., a (0,+1) event window, and three days, i.e., a (−1,+1) event window. The results from these estimations are consistent with the findings reported in Table 8, although the significance levels on the nonaudit fee variables decline as the event window is expanded. For the two-day and three-day event windows, FEERATIO is significant at the 0.06 and 0.21 levels, RANKNON is significant at the 0.17 and 0.52 levels, and FEERESIDUAL is significant at the 0.11 and 0.25 levels.

Taken together, the results suggest a negative association between nonaudit fees and abnormal returns on the date the fees were disclosed, but no association between audit fees and abnormal returns. However, the fee variables explain only 1 percent of the cross-sectional variation in abnormal returns, and the magnitude of the coefficient estimates implies a small economic effect. For example, the coefficient estimate of −0.02 on FEERESIDUAL multiplied by the interquartile range of this variable (0.28) suggests that an increase in unexpected nonaudit fees from the lowest to highest quartile is expected to decrease share values by approximately 0.50 percent. A final caveat to this analysis is that the disclosure of unexpectedly high nonaudit fees may signal other value-relevant information to the market, such as poor management quality or operational difficulties. However, 99 percent of the sample released fiscal 2000 earnings prior to filing their proxy statements, thus pre-empting at least a portion of the performance-related information content of the fee data.

V. SUMMARY AND CONCLUSION

This paper provides evidence on the association between auditor fees and earnings management, and between auditor fees and share values on the date the fees are disclosed. We find a positive association between nonaudit fees and the likelihood of reporting a small earnings surprise, the magnitude of absolute discretionary accruals, and the magnitudes of

---

[22] Overlapping event dates can produce dependence across firm-events. To control for the effects of cross-sectional dependence on significance levels, we create distributions of regression coefficients by estimating the model on 200 nonevent dates beginning at date t−141, where t is the event date. Nonevent dates are taken from the intervals [t−141, t−2] and [t+2, t+61]. Because all observations are shifted to nonevent dates as a block, observations in each nonevent date regression are temporally related as they are in the event date regression. Bootstrap p-values determined using this method are qualitatively similar to the results reported in Table 8.

TABLE 8
Abnormal Returns to the Disclosure of Auditor Fee Information

$$ARET = \beta_0 + \beta_1 FEEVAR + \beta_2 LOGMVE + \varepsilon$$

Variable	*Coeff. Est.* *(p-value)*	*Coeff. Est.* *(p-value)*	*Coeff. Est.* *(p-value)*	*Coeff. Est.* *(p-value)*	*Coeff. Est.* *(p-value)*	*Coeff. Est.* *(p-value)*	*Coeff. Est.* *(p-value)*	*Coeff. Est.* *(p-value)*
Intercept	0.01 (0.03)	0.01 (0.70)	0.01 (0.01)	−0.01 (0.95)	0.01 (0.01)	0.01 (0.88)	0.01 (0.43)	−0.01 (0.38)
FEERATIO	−0.01 (0.10)	−0.02 (0.01)						
RANKNON			−0.01 (0.09)	−0.01 (0.01)				
RANKAUD			0.01 (0.81)	−0.01 (0.81)				
RANKTOT					−0.01 (0.05)	−0.01 (<0.01)		
FEERESIDUAL							−0.02 (0.01)	−0.02 (0.01)
LOGMVE		0.01 (0.03)		0.01 (<0.01)		0.01 (<0.01)		0.01 (0.20)
R^2	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
Nobs.	1,811		1,811		1,811		1,693	

ARET = the market model prediction error on the proxy filing date, estimated using firm-level regressions of raw returns on the corresponding NYSE size-decile portfolio returns over the 100-day period ending 30 days prior to the proxy filing date; and

FEERESIDUAL = the prediction error from the model estimated in the Appendix.

All other variables are defined in Table 4, where FEEVAR indicates the alternative specifications of the fee variables, (1) FEERATIO, (2) RANKNON and RANKAUD, (3) RANKTOT, or (4) FEERESIDUAL.

both income-increasing and income-decreasing discretionary accruals. There is a negative association between audit fees and these earnings management indicators. These results are robust to alternative variable definitions and model specifications. In contrast, we find no association between total fees and any of these earnings management indicators, indicating that combining audit and nonaudit fees into a single measure masks their differential incentive effects. We also find no association between nonaudit fees and the likelihood of reporting a small increase in earnings, although audit fees are negatively associated with this earnings management indicator. Evidence from auditor-specific regressions indicates that the results are not driven by any particular auditor. Finally, we find evidence of a negative association between the disclosure of nonaudit fees and share values, although the reaction is small in economic terms and insignificant when measured over longer event windows.

These results suggest a number of opportunities for further analysis. For example, our tests indicate that audit and nonaudit fees create different incentive effects, despite arguments to the contrary in prior research (e.g., Hansen and Watts 1997; Reynolds and Francis 2000). Analytical and empirical research exploring this finding would contribute to our knowledge of auditor-client contracting. Our tests also suggest that size is an important conditioning variable in explaining the incentive effects of audit and nonaudit fees. Future research could consider this and other economic circumstances that affect auditor incentives. An exploration of alternative reporting contexts, e.g., audit opinions (see DeFond et al. 2002), earnings restatements, or financial statement fraud, would provide a more complete understanding of the financial statement consequences of auditors' provision of nonaudit services. Finally, our analyses are based on the first year of disclosed fees. As additional years of data become available, research on changes in the market for audit and nonaudit services and the effect of these changes on auditor incentives and financial reporting quality would be of interest.

APPENDIX
ESTIMATION OF UNEXPECTED NONAUDIT FEES

In this appendix, we develop a model to estimate the unexpected portion of nonaudit fees. We base our model on results in Parkash and Venable (1993) and Firth (1997) who examine the decision to purchase nonaudit services from the auditor, conditional on purchases of auditing services. They argue that a company facing high agency costs demands high-quality audits and restricts the purchase of other services from their auditor.

We estimate the following model using all observations with the necessary data:

$$\begin{aligned}\text{FEERATIO} = {} & \sum_{i=1}^{13} \alpha_1\text{IND}_i + \beta_1\text{BIGFIVE} + \beta_2\text{AUDTEN} + \beta_3\text{ROA} + \beta_4\text{LOSS} \\ & + \beta_5\text{ANNRET} + \beta_6\text{\%INST} + \beta_7\text{CFO} + \beta_8\text{LEVERAGE} \qquad \text{(A1)} \\ & + \beta_9\text{INVREC} + \beta_{10}\text{FIN/ACQ} + \beta_{11}\text{LOGMVE} + \beta_{12}\text{M/B} + \varepsilon\end{aligned}$$

We include industry intercepts, IND_i, for each of the i = 1 to 13 industries identified in Panel C of Table 1. An indicator variable (BIGFIVE) controls for the possibility that Big 5 auditors are able to provide more nonaudit services and/or charge a premium for their services. AUDTEN, the number of years the auditor has audited the financial statements of the company, controls for the length of the auditor-client relationship.

Performance may proxy for agency costs as well as for the potential benefits from purchasing outside consulting. We include two measures of accounting performance, ROA,

TABLE A1
Estimation of Unexpected Nonaudit Fees

$$
\begin{aligned}
FEERATIO = \sum_{i=1}^{13} \alpha_i IND_i &+ \beta_1 BIGFIVE + \beta_2 AUDTEN + \beta_3 ROA + \beta_4 LOSS \\
&+ \beta_5 ANNRET + \beta_6 \%INST + \beta_7 CFO + \beta_8 LEVERAGE \\
&+ \beta_9 INVREC + \beta_{10} FIN/ACQ + \beta_{11} LOGMVE + \beta_{12} M/B + \varepsilon
\end{aligned}
$$

Variable	*Coefficient Estimate*	*t-statistic*	*p-value*
Agriculture	−0.10	−1.29	0.20
Mining and construction	−0.03	−1.09	0.27
Food	−0.07	−2.20	0.03
Textiles and printing/publishing	−0.02	−0.94	0.34
Chemicals	−0.02	−0.72	0.47
Pharmaceuticals	−0.01	−0.37	0.71
Extractive	−0.10	−3.79	<0.01
Durable manufacturers	0.01	−0.30	0.76
Transportation	0.01	0.74	0.46
Utilities	−0.03	−1.62	0.11
Retail	−0.07	−2.40	0.02
Services	−0.03	−1.77	0.08
Computers	0.14	6.08	<0.01
BIGFIVE	0.07	4.46	<0.01
AUDTEN	−0.01	−3.62	<0.01
ROA	−0.04	−2.74	<0.01
LOSS	0.01	1.32	0.19
ANNRET	−0.03	−4.69	<0.01
%INST	−0.01	−0.63	0.53
CFO	0.02	1.16	0.25
LEVERAGE	0.02	0.99	0.32
INVREC	−0.03	−1.08	0.28
FIN/ACQ	0.07	8.08	<0.01
LOGMVE	0.05	20.01	<0.01
M/B	0.01	−5.87	<0.01
Adjusted R^2	0.31		
n	2,520		

(*Continued on next page*)

TABLE A1 (Continued)

Industry membership is determined by SIC code as follows: agriculture (0100–0999), mining and construction (1000–1999, excluding 1300–1399), food (2000–2111), textiles and printing/publishing (2200–2799), chemicals (2800–2824, 2840–2899), pharmaceuticals (2830–2836), extractive (1300–1399, 2900–2999), durable manufacturers (3000–3999, excluding 3570–3579 and 3670–3679), transportation (4000–4899), utilities (4900–4999), retail (5000–5999), services (7000–8999, excluding 7370–7379), and computers (3570–3579, 3670–3679, 7370–7379).

BIGFIVE = 1 if the firm's auditor is a Big 5 firm, and 0 otherwise;
AUDTEN = number of years that the auditor has audited the firm's financial statements;
ROA = net income divided by average total assets;
LOSS = 1 if the firm reported a net loss in fiscal 2000, and 0 otherwise;
CFO = cash from operations, deflated by average total assets;
ANNRET = percentage compounded monthly stock return for calendar 2000 minus the CRSP value-weighted market index;
%INST = percent of shares held by institutions (reported by Spectrum);
LEVERAGE = ratio of total liabilities to total assets;
INVREC = inventory plus accounts receivable as a percentage of total assets;
FIN/ACQ = 1 if the firm issued securities or acquired another company in fiscal 2000, and 0 otherwise;
LOGMVE = log of market value of equity; and
M/B = market-to-book ratio.

defined as net income divided by average total assets, and LOSS, an indicator variable equal to 1 if the firm reports a net loss. We also include stock return performance, ANNRET, the compounded CRSP monthly return in calendar 2000. Institutional ownership (%INST) is the percentage of shares held by institutions (reported by Spectrum), and LEVERAGE is the ratio of total liabilities to total assets.

We include operating cash flow scaled by average total assets (CFO) to control for the ability to purchase outside consulting services. In addition, we control for the extent and complexity of audit services by including INVREC, equal to inventory plus accounts receivable as a percent of total assets. Rapidly growing companies may demand more consulting services, and thus we include FIN/ACQ, an indicator variable for firms that issued securities or acquired another company in 2000, and the market to book ratio (M/B). Finally, we control for firm size with the log of market value of equity (LOGMVE).

Summary statistics from the estimation of Equation (A1) are reported in Table A1. The model explains 31 percent of the cross-sectional variation in FEERATIO, comparable to the explanatory power of the models in Parkash and Venable (1993) and Firth (1997). The results indicate that firms in the food, extractive, and retail industries purchase fewer nonaudit services, while firms in the computer industry purchase more nonaudit services. Consistent with Parkash and Venable (1993) and Firth (1997), we find that firms with Big 5 auditors and poorly performing firms, as measured by ROA and ANNRET, purchase more nonaudit services. However, neither institutional ownership nor leverage is significant. Acquisition and financing activities also contribute to significantly higher nonaudit fees. Finally, we find that nonaudit fees are increasing in firm size and decreasing in auditor tenure and growth.

REFERENCES

Abbott, L., S. Parker, G. Peter, and D. V. Rama. 2001. Audit, nonaudit and information technology fees: Some empirical evidence. Working paper, University of Memphis.

Abdel-khalik, A. R. 1990. The jointness of audit fees and demand for MAS: A self-selection analysis. *Contemporary Accounting Research* 6: 295–322.

Antle, R., P. A. Griffin, D. J. Teece, and O. E. Williamson. 1997. *An Economic Analysis of Auditor Independence for a Multi-Client, Multi-Service Public Accounting Firm.* Report prepared for the AICPA by The Law & Economics Consulting Group, Inc. Stamford, CT: American Institute of Certified Public Accountants.

Ashbaugh, H., R. LaFond, and B. W. Mayhew. 2002. Do nonaudit services compromise independence? Further evidence. Working paper, University of Wisconsin–Madison.

Arruñada, B. 1999. *The Economics of Audit Quality: Private Incentives and the Regulation of Audit and Nonaudit Services.* Norwell, MA: Kluwer Academic Publishers.

Barkess, L., and R. Simnett. 1994. The provision of other services by auditors: Independence and pricing issues. *Accounting and Business Research* 24: 99–108.

Bartov, E., F. A. Gul, and J. S. L. Tsui. 2001. Discretionary accruals models and audit qualifications. *Journal of Accounting and Economics* 30: 421–452.

Beck, P. J., T. J. Frecka, and I. Solomon. 1988a. A model of the market for MAS and audit services: Knowledge spillovers and auditor-auditee bonding. *Journal of Accounting Literature* 7: 50–64.

———, ———, and ———. 1988b. An empirical analysis of the relationship between MAS involvement and auditor tenure: Implications for auditor independence. *Journal of Accounting Literature* 7: 65–84.

Becker, C. L., M. L. DeFond, J. Jiambalvo, and K. R. Subramanyam. 1998. The effect of audit quality on earnings management. *Contemporary Accounting Research* 15: 1–24.

Beeler, J. D., and J. E. Hunton. 2001. Contingent economic rents: Precursors to predecisional distortion of client information. Working paper, University of South Carolina.

Brown, L. D. 2001. A temporal analysis of earnings surprises: Profits versus losses. *Journal of Accounting Research* 39: 221–241.

Burgstahler, D., and I. Dichev. 1997. Earnings management to avoid earnings increases and losses. *Journal of Accounting and Economics* 24: 99–126.

———, and M. Eames. 1998. Management of earnings and analyst forecasts. Working paper, University of Washington.

Chung, H., and S. Kallapur. 2001. Client importance, nonaudit services, and abnormal accruals. Working paper, Purdue University.

Coffee, J. C. 2001. The acquiescent gatekeeper: Reputational intermediaries, auditor independence, and the governance of accounting. Working paper, Columbia University School of Law.

Craswell, A. T., J. R. Francis, and S. L. Taylor. 1995. Auditor brand name reputations and industry specializations. *Journal of Accounting and Economics* 20: 297–322.

Davis, L. R., D. N. Ricchiute, and G. Trompeter. 1993. Audit effort, audit fees, and the provision of nonaudit services to audit clients. *The Accounting Review* 68: 135–150.

DeAngelo, H., L. DeAngelo, and D. J. Skinner. 1994. Accounting choice in troubled companies. *Journal of Accounting and Economics* 10: 193–225.

DeAngelo, L. 1981. Auditor size and audit quality. *Journal of Accounting and Economics* 3: 183–199.

Dechow, P. M., R. G. Sloan, and A. P. Sweeney. 1995. Detecting earnings management. *The Accounting Review* 70: 193–225.

———, ———, and ———. 1996. Causes and consequences of earnings manipulation: An empirical analysis of firms subject to enforcement actions by the SEC. *Contemporary Accounting Research* 13: 1–36.

DeFond, M. L., and J. Jiambalvo. 1991. Incidence and circumstances of accounting errors. *The Accounting Review* 66: 643–655.

———, and ———. 1993. Factors related to auditor-client disagreements over income-increasing accounting methods. *Contemporary Accounting Research* 9: 411–431.

———, and ———. 1994. Debt covenant violation and manipulation of accruals. *Journal of Accounting and Economics* 17: 145–176.

———, K. Raghunandan, and K. R. Subramanyam. 2002. Do nonaudit service fees impair auditor independence? Evidence from going concern opinions. Working paper, University of Southern California.

Degeorge, F., J. Patel, and R. Zeckhauser. 1999. Earnings management to exceed thresholds. *Journal of Business* 72: 1–34.

Dopuch, N., R. King, and R. Schwartz. 2001. Auditors' independence: Appearance vs. fact. Working paper, Washington University.

Feroz, E. H., K. Park, and V. S. Pastena. 1991. The financial and market effects of the SEC's accounting and auditing enforcement releases. *Journal of Accounting Research* 29: 107–142.

Firth, M. 1990. Auditor reputation: The impact of critical reports issued by government inspectors. *RAND Journal of Economics* 21: 374–387.

———. 1997. The provision of nonaudit services by accounting firms to their audit clients. *Contemporary Accounting Research* 14: 1–21.

Francis, J. D., D. Philbrick, and K. Schipper. 1994. Shareholder litigation and corporate disclosures. *Journal of Accounting Research* 32: 137–164.

Francis, J. R., E. L. Maydew, and H. C. Sparks. 1999. The role of Big 6 auditors in the credible reporting of accruals. *Auditing: A Journal of Practice & Theory* 18: 17–34.

———, and B. Ke. 2002. Do nonaudit services compromise auditor independence? Working paper, University of Missouri–Columbia.

Glezen, G. W., and J. A. Millar. 1985. An empirical investigation of stockholder reaction to disclosures required by ASR No. 250. *Journal of Accounting Research* 23: 859–870.

Gore, P., P. F. Pope, and A. K. Singh. 2001. Nonaudit services, auditor independence, and earnings management. Working paper, Lancaster University.

Hansen, S. C., and J. S. Watts. 1997. Two models of the auditor-client interaction: Tests with United Kingdom data. *Contemporary Accounting Research* 14: 23–50.

Johnston, J. 1984. *Econometric Methods*. 3rd edition. New York, NY: McGraw-Hill, Inc.

Jones, J. J. 1991. Earnings management during import relief investigations. *Journal of Accounting Research* 29: 193–228.

Kothari, S. P., A. L. Leone, and C. E. Wasley. 2002. Performance matched discretionary accrual measures. Working paper, University of Rochester.

Lys, T., and R. L. Watts. 1994. Lawsuits against auditors. *Journal of Accounting Research* 32: 65–93.

Magee, R., and M. Tseng. 1990. Audit pricing and independence. *The Accounting Review* 65: 315–336.

Matsumoto, D. 2002. Management's incentives to avoid negative earnings surprises. *The Accounting Review* 77 (3): 483–514.

McNichols, M. F. 2000. Research design issues in earnings management studies. *Journal of Accounting and Public Policy* 19: 313–345.

Nelson, M. W., J. A. Elliott, and R. L. Tarpley. 2002. Evidence from auditors about managers' and auditors' earnings-management decisions. *The Accounting Review* (Supplement): 175–202.

Palmrose, Z-V. 1986. The effect of nonaudit services on the pricing of audit services: Further evidence. *Journal of Accounting Research* 24: 405–411.

Parkash, M., and C. F. Venable. 1993. Auditee incentives for auditor independence: The case of nonaudit services. *The Accounting Review* 68: 113–133.

Public Oversight Board (POB). 2000. *Panel on Audit Effectiveness: Report and Recommendations*. Stamford, CT: American Institute of Certified Public Accountants.

Reynolds, J. K., and J. R. Francis. 2000. Does size matter? The influence of large clients on office-level auditor reporting decisions. *Journal of Accounting and Economics* 30: 375–400.

Scheiner, J. H. 1984. The impact of SEC nonaudit service disclosure requirements on independent auditors and their clients: An empirical investigation. *Journal of Accounting Research* 22: 789–797.

Securities and Exchange Commission (SEC). 2000. *Final Rule: Revision of the Commission's Auditor Independence Requirements*. Washington, D.C.: Government Printing Office.

Simunic, D. 1984. Auditing, consulting, and auditor independence. *Journal of Accounting Research* 22: 679–702.

Teoh, S. H., and T. J. Wong. 1993. Perceived auditor quality and the earnings response coefficient. *The Accounting Review* 68: 346–367.

Trompeter, G. 1994. The effect of partner compensation schemes and generally accepted accounting principles on audit partner judgment. *Auditing: A Journal of Practice & Theory* 13: 56–69.

Warfield, T. D., J. J. Wild, K. L. Wild. 1995. Managerial ownership, accounting choices, and informativeness of earnings. *Journal of Accounting and Economics* 20: 61–91.

Weil, J., and J. Tannenbaum. 2001. Big companies pay audit firms more for other services. *Wall Street Journal* (April 10): C1.

White, H. 1980. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica* 48: 817–838.

THE ACCOUNTING REVIEW
Vol. 77
Supplement 2002
pp. 107–114

DISCUSSION OF The Relation between Auditors' Fees for Nonaudit Services and Earnings Management

William R. Kinney, Jr.
The University of Texas at Austin
Robert Libby
Cornell University

I. INTRODUCTION

Frankel et al. (2002) (hereafter FJN) present a timely paper using a new data set to test several propositions, including one suggested by the Securities and Exchange Commission (SEC) about the relation between nonaudit fees paid by a registrant to its auditor and the registrant's earnings quality. In this discussion we use comments from participants at *The Accounting Review* Conference on Quality of Earnings and our own analyses to evaluate FJN's approach from both conceptual and operational perspectives.

We organize our discussion around a "Predictive Validity Model" (displayed later in Figure 1) that shows conceptual relations between independent and dependent variables, as well as their operational measures and treatment of other causal factors. We use this model to describe and evaluate the authors' approach, which blends complex behavioral and contractual relations with aggregate observable fee, accounting, and market price data. We hope that the discussion will encourage others to conduct additional research to address the important questions raised by both FJN and the SEC.

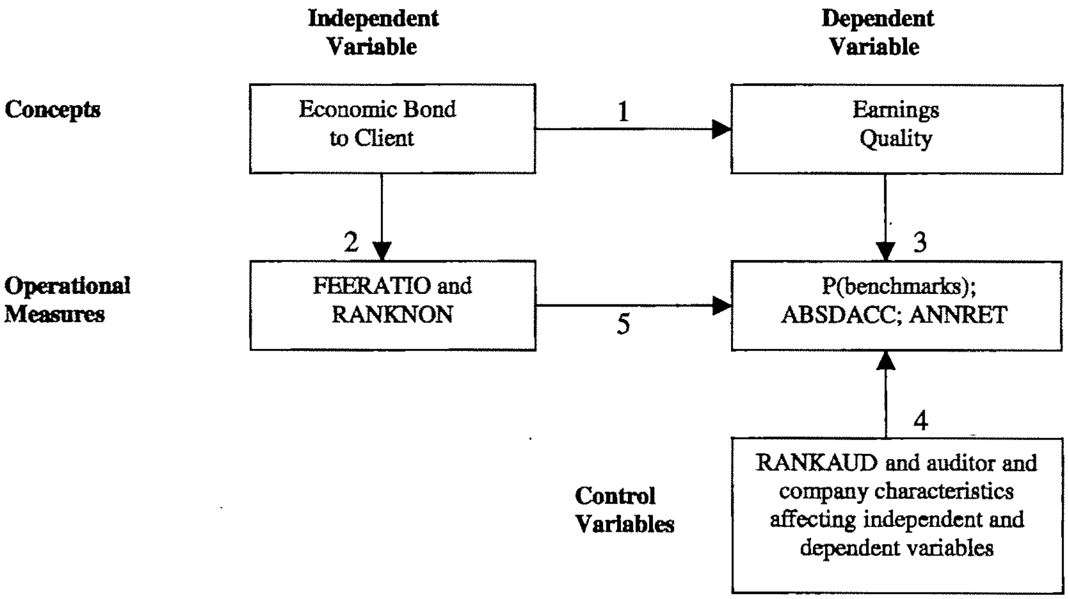
II. FJN's APPROACH

The underlying proposition examined by FJN is that the strength of the audit firm's economic dependence on the client, or *economic bond*, consciously or unconsciously reduces the auditor's *independence*, or willingness to resist client-induced biases in the reporting process. The lack of auditor independence results in lower *earnings quality*. The underlying cause (the economic bond) and the effect (lowered earnings quality) are the two theoretical constructs examined. Both are defined conceptually, and are not directly observable. Their relationship forms Link 1 in the first (or concepts) row of the Predictive Validity Model shown in Figure 1.[1]

---

[1] The Predictive Validity Model is based on Runkel and McGrath (1972) and Libby (1981), who used the model to analyze accounting research questions.

FIGURE 1
Predictive Validity Model, as Applied to Auditor-Client Relations by FJN



FEERATIO = the ratio of financial information system and other fees to total fees;
RANKNON = the percentile rank of nonaudit fees, by auditor;
the benchmark SURPRISE = 1 if the firm meets or just beats the consensus analyst forecast (i.e., forecast error of 0 ¢ or 1¢) for fiscal 2000, and 0 otherwise;
the benchmark INCREASE = 1 if the firm reports a small increase in earnings relative to fiscal 1999, and 0 otherwise;
ABSDACC = the absolute value of total accruals, equal to net income minus cash from operations, deflated by average total assets; and
ANNRET = the percentage compounded monthly stock return for calendar 2000 minus the CRSP value-weighted market index.

At the operational level, FJN first adopt the SEC's proposition that the relevant operational measure of the auditor's economic bond to the client is the ratio of nonaudit fees to total fees received from a client (denoted FEERATIO). They also measure economic bond as the rank of nonaudit fees across clients for an audit firm in their sample (denoted RANKNON), assuming that the rank of the magnitude of nonaudit fees (irrespective of the magnitude of audit fees) compromises auditor independence. The rank of audit fees (RANKAUD) is also tracked, but is not presumed to affect auditor independence. The relation between the concept economic bond and its operational measures is Link 2 in the model.

The authors focus on one element of earnings quality, "management's deliberate intervention in the earnings process in the form of earnings management," and use three types of surrogate measures of earnings management: (1) P(benchmark) defined as the proportion of clients just meeting or beating earnings benchmarks (consensus analyst forecasts and prior period earnings), (2) ABSDACC or absolute discretionary accruals, and (3) ARET, the market price reaction to disclosure of fees, which implies a price reduction for presumed earnings management. These variables comprise the operational dependent variables. The relation between earnings quality and its three measures is Link 3 in the Predictive Validity Model. Other auditor and company characteristics comprise control variables whose relation to the operational dependent variables is Link 4 in the model.

As in all empirical studies, FJN draw inferences about Link 1, the association between concepts, by assessing Link 5, the association between operational measures of those concepts. The validity of this inference depends on the validity of Links 2 and 3, the operationalization of the concepts, and whether the effects in Link 4 are adequately controlled. Thus, most concerns about FJN's inferences, and directions for future research, can be classified by their relation to:

- Link 1—inconsistency in theory.
- Links 2 and 3—faulty operationalization.
- Link 4—failure to control for other potential causes.
- Link 5—data or statistical problems.

In the discussion to follow, we categorize our comments and those of the Conference participants in terms of the five links in Figure 1.

Link 1: Economic Bond and Earnings Quality

FJN propose that a stronger economic bond between audit/consulting firm and client will reduce the quality of reported earnings, other things equal. Auditor independence impairment, or reduced willingness to resist client-induced biases in reported earnings, is the presumed intervening process through which economic bond affects earnings quality, but is not measured in this study. As a consequence, much of FJN's discussion of SEC independence rules is not directly relevant to the issues examined in the paper.[2]

Better conceptual definitions can improve measurement of the concepts in all empirical work. We identify three issues related to the definition of economic bond that deserve increased attention in future research. First, FJN do not appear to consider whether high audit fees might also increase the economic bond. However, because the authors present no reason to distinguish between audit and nonaudit fees (or profits) as determinants of economic bond, others might reasonably predict that operational measures of economic bond such as nonaudit fees, audit fees, and their sum would *all* be correlated positively with operational measures of earnings quality. This is a serious conceptual omission.

Second, the concept of economic bond could be refined to distinguish between *total* fees (or profits) and fees in excess of those expected from observable firm circumstances. As the authors suggest, both acquisitions and complex tax environments increase expected demands for both audit and consulting services. More insidious effects on economic bond may result from *unexpected* nonaudit and audit fees that may more accurately be likened to attempted bribes. FJN's estimation of unexpected nonaudit fees in the Appendix suggests they have considered this issue, but the same approach was not applied to audit fees and total fees, and the associations between unexpected fees and the earnings management measures were not proposed or reported. Unexpected fees may also better capture the profitability of the services provided.

The benefits to building better models of expected audit fees can be seen in an analysis of Enron's fee disclosures (audit fees of $25 million and nonaudit fees of $27 million).

---

[2] In the final section, we discuss, as a suggestion for future research, a model that makes explicit some components implicit in the FJN independence argument.

Enron's resulting FEERATIO of .52 is close to the median (.51) reported by FJN in Table 2, but the audit fee is more than 250 percent of the expected fee computed using a 1997 survey of audit fees as a proportion of client revenue.[3] Thus, while Enron would be rated as about average in economic bond based on FEERATIO, the unexpected audit fee may imply a much stronger economic bond, one that would remain even in the absence of nonaudit fees. The result of an expected fees exercise would be estimates of the unexpected or "excess" profitability of an audit client, irrespective of whether the excess results from a particular service or client-billing peculiarities of the audit firm.

The third conceptual issue related to economic bond that should be reconsidered is the form of the *expected* relationship between audit and nonaudit fees as components of the economic bond. As the results presented in FJN suggest, this relationship could be complex. For example, a willingness to lower audit fees to maximize potential total client fees or profits could cause unusual conditional relationships between the two fee components and economic bond. Other conceptual interactions probably exist.

Concerning the conceptual definition of earnings quality, we agree with the authors' choice of the most often studied element of earnings quality, earnings management. This choice allows the authors to compare their results with existing studies. Future research can broaden the number of earnings quality dimensions analyzed.

Link 2: Economic Bond and Nonaudit Fee Measures

Because FJN do not distinguish conceptually between audit fee and nonaudit fee sources of economic bond, both their interpretation of FEERATIO and their failure to include RANKNON and RANKAUD as components of economic bond are problematic. Since FEERATIO increases when audit fees decrease and RANKAUD is negatively related to earnings management, the results for FEERATIO, and for RANKNON and RANKAUD, are internally consistent (and consistent with prior research). They are inconsistent, however, with FJN's theory if the economic loss (bond) resulting from a client departure would be the total fees paid to the auditor. The lack of association between total fees and earnings management proxies adds to questions about the interpretability of the results.

Conference participants suggested a number of possible avenues for investigating these theory-inconsistent results. Some comments focused on other concepts for which RANKAUD may be proxying, including size[4] and audit fee low-balling. If larger clients (which have larger audit fees) are less likely to manage earnings, then audit fees may be negatively associated with earnings management. Similarly, a negative association could also result if low-balling audit fees induces more auditor acceptance of earnings management attempts (Beeler and Hunton 2001).

Participants also suggested that future studies examine the role of nonaudit fees in audit partner compensation and whether the economic bond operates at the office level, as opposed to the firm level implicit in the RANKAUD and RANKNON variables. For example, a client with a low value for RANKNON or RANKAUD may represent a huge proportion of fees for a small office of an audit firm.

A number of questions were raised about the SEC's categorization of fees into the audit and nonaudit categories. It was suggested that, because 2000 was the first year for fee disclosure, the SEC categories might not be well understood by all registrants, leading to differences in their application across registrants both within a particular audit firm's clients and perhaps across audit firms (although at least one audit firm provided detailed instructions on how to break down nonaudit fees). These issues argue for an analysis by audit firm, including models that allow slopes to differ by audit firm as well as intercepts (now included by FJN in Table 7).

---

[3] Based on a survey of 340 members of Manufacturers Alliance (Arlington, VA) reported in *Public Accounting Report*, November 30, 1997, the audit fee for firms with $20 billion or more in revenues was $92 per million in revenues. Enron's $100 billion in revenues would imply an audit fee of $9,200,000.

[4] The proxy used for size may not be fully effective.

Measurement concerns were also raised about items normally related to audit services (e.g., pension plan audits, consultation on accounting standards or transactions, and tax work) that are included in the SEC definition of nonaudit fees. In addition, comparisons of U.S. results with studies in other countries will require reconciliation of the differences in definitions of nonaudit fees across countries. Finally, interpretation of results may be impaired by the use of a single year (the first year the disclosures were required, and concentration of almost 60 percent of the proxies filed in April 2001) that may be idiosyncratic with respect to the measured relations. While each of these concerns may add error to the analyses, no Conference participant suggested specific directional biases that may have been induced.

Link 3: Earnings Quality and Empirical Measures

FJN measure earnings management using two measures popular in archival studies: the probability of meeting or beating benchmarks (P(benchmark)), and the absolute value of discretionary accruals (ABSDACC). A strength of the paper is the fact that these two different types of measures provide some convergent results, reducing the possibility that the results are due to method bias.

Most Conference participants' concerns about measurement of earnings quality pointed to inherent contradictions in the earnings management literature, as opposed to specific problems in FJN's study. For example, the benchmark measure categorizes all firms that meet or beat the consensus forecast as having low-quality earnings, regardless of whether the goal was achieved through earnings management, expectations management, reduction of uncertainty, or improvements in operations. Conversely, a firm that is consistently well below or well above the consensus forecast will be deemed to have high-quality earnings by this criterion, even though recorded results may be far from the consensus due to a weak accounting system or attempts to manipulate earnings by large amounts. Similarly, discretionary accruals categorize, as earnings managers, firms that engage in transactions that take a long time to complete, or that involve complex judgments and estimates.

Finally, short-window price reactions to the release of proxy statements confound the release of absolute and relative fee information with other conditions or events noted in the proxy statement as well as other contemporaneous events occurring on the relatively few proxy filing dates. Also, high nonaudit fees may cause a negative price reaction because of the inference that the company "needed" a lot of consulting, instead of an inference that the audit and the resulting earnings were of low quality.

Link 4: Control Variables

Rather than using a system of equations (one to model fees, and one to model surrogates for earnings quality), FJN use a single equation simultaneously to model earnings quality surrogates and to control for nonaudit and audit fee determinants. They include a variety of control variables in the earnings quality regressions. Some of the control variables are alleged to decrease the operational dependent variable, earnings quality (litigation risk, growth, institutional ownership, and a reported profit vs. loss). Other included control variables are alleged to increase the operational independent variable, purchases of nonaudit services (acquisition and financing activity and poor performance). Still others measure additional factors that increase the conceptual independent variable, economic bond (longer auditor tenure and smaller auditor size). Firms with higher cash flows from operations are seen as less likely to manage earnings *and* more able to afford nonaudit services. So this variable is thought to be a determinant of *both* the operational dependent and independent variables. And, of course, client size is included as a control because it is related to most things.

Controlling for determinants of the operational independent (fee) variables in a single equation modeling earnings management makes the interpretation of the coefficients on the fee variables problematic. Including these variables could hide, accentuate, or even reverse the sign of associations between the fee variables and the earnings management variables. At the least, their inclusion eliminates sources of variation in nonaudit and audit fees that could affect earnings management.

Link 5: Nonaudit Fees and Earnings Quality Measures

Participants requested additional analyses to allow some assessment of the stability of the basic results, first among these being the correlation matrix for all of the variables (now added in FJN's Table 4). Second, in addition to the regressions provided, Conference participants suggested including the regressions with only the controls for the dependent variable followed by a regression including these same controls and the fee variable(s), to judge the marginal predictive power of the fee variables (analyses are now added to FJN's Tables 5–8).

As to the single equation approach to modeling, participants suggested separate but related models of determinants of both fees and earnings management in a system of equations. One approach was to include exogenous variables suspected to affect the independent variable nonaudit fees in a two-stage regression, using the predicted value of the nonaudit fee variable as an instrumental variable in the second stage. These same types of analyses were suggested for the total fee measure, RANKTOT, to probe the apparent lack of association between this measure of economic bond and earnings management. Other participants suggested a simultaneous-equations approach to account for the independent and dependent variables whose values may be jointly determined.

Participants also raised completeness issues. Several suggested incorporating as control variables corporate governance factors previously found to be related to earnings management. In addition, omissions in the discussion of the implications of certain results were noted, the most important being the lack of association of earnings management with RANKTOT and the negative association with RANKAUD. Several participants wanted to know why the economic bond argument should apply only to RANKNON and not to RANKTOT, and perhaps apply negatively to RANKAUD. Also noted was the lack of explanation for the disappearance of the nonaudit fee relation for larger firms. Is this an anomaly, or does the economic bond theory predict a conditional size relation? Several participants also questioned FJN's emphasis on the significant findings using ABSDACC and the earning forecast benchmark, and the lack of emphasis on the nonsignificant findings using the earnings change benchmark, where, of the fee measures, only the RANKAUD variable remain significant.

FJN conduct three tests of the stock price effects of public announcement of fees via registrant proxy filings. The authors take a short-window events study approach at the announcement date and predict that stock prices will react negatively to high, and higher than expected, nonaudit fees. The coefficients for the FEERATIO, and the RANKNON and RANKAUD models are insignificantly different from zero, suggesting no measurable effects. R^2s for the two models are .01. The third model, using a within-sample calculation of unexpected nonaudit fees (FEERESIDUAL), also has an R^2 of .01, but the coefficient is significantly negative.

FJN interpret their overall stock price reaction results as "modest evidence consistent with investors reducing valuations as compensation for a perceived reduction in audit quality." Taken at face value, some participants wondered why *unexpected* FEERESIDUAL should compromise audit quality and lead to price declines while the nonaudit fee and audit fee measured by FEERATIO and both RANKNON and RANKAUD should not.

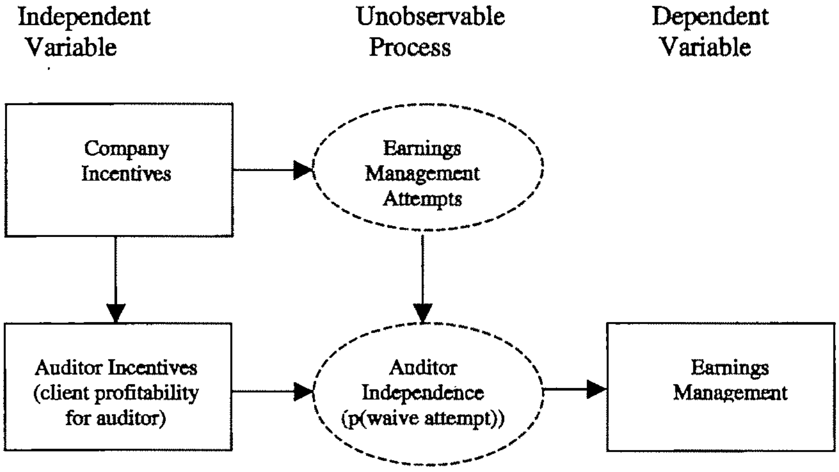
III. CONCLUSIONS

Overall, FJN have presented an ambitious analysis of a new data set and have made an important first pass at addressing some issues of great interest to regulators, practitioners, and scholars. Their data gathering was as comprehensive as was possible at the time, and their analysis provides documentation of empirical relations that suggest a basis for further investigation. The use of multiple measures helps in understanding possible causal links related to the constructs of economic bond and earnings quality. At the same time, however, various research design and measurement issues discussed above make their results difficult to interpret.

Issues related to the conceptual definition of economic bond cause the greatest problems. First among these is the lack of association between RANKTOT (the rank of total fees within each audit firm) and the earnings management measures. Second is the statistically negative relation between RANKAUD and the earnings management measures. Higher fees of *either* kind would presumably increase the economic bond. The authors' explanation that the negative relation to audit fee has been found in other studies does not explain why the relation *should* be negative in the present circumstance, nor help with the interpretation of its implications for their economic bond theory.

Finally, we suggest that the conceptual relations between management and auditor incentives and earnings management are more complex than suggested in FJN. The model

FIGURE 2
Conceptual Determinants of Earnings Management


Independent Variable
Unobservable Process
Dependent Variable
Company Incentives
Earnings Management Attempts
Auditor Incentives (client profitability for auditor)
Auditor Independence (p(waive attempt))
Earnings Management

in Figure 2 recognizes that management incentives affect both the magnitude of earnings management attempts and auditor incentives to resist those attempts (Nelson et al. 2002).

The model points out that an empirical association between fees and earnings management could result from firms with higher fees *attempting* more earnings management and/or from reduced independence, neither of which is observable in archival studies. This suggests that careful modeling of both management and auditor incentives, combined with careful estimation of expected audit and nonaudit fees, may provide useful directions for future research.

REFERENCES

Beeler, J. D., and J. E. Hunton. 2001. Contingent economic rents: Precursors to predecisional distortion of client information. Working Paper, University of South Florida.

Frankel, R. M., M. F. Johnson, and K. K. Nelson. 2002. The relation between auditors' fees for nonaudit services and earnings management. *The Accounting Review* (Supplement): 71–106.

Libby, R. 1981. *Accounting and Human Information Processing: Theory and Applications*. Englewood Cliffs, NJ: Prentice Hall, Inc.

Nelson, M., J. Elliott, and R. Tarpley. 2002. Evidence from auditors about managers' and auditors' earnings-management decisions. *The Accounting Review* (Supplement): 175–202.

Runkel, P., and J. McGrath. 1972. *Research on Human Behavior–A Systematic Guide to Method*. New York, NY: Holt, Rinehart, and Winston, Inc.

THE ACCOUNTING REVIEW
Vol. 77
Supplement 2002
pp. 115–133

Empirical Evidence on the Evolution of International Earnings

Judy Land
Mark H. Lang
The University of North Carolina at Chapel Hill

ABSTRACT: We examine whether cross-country differences in earnings-to-price multiples have changed between 1987–1992 and 1994–1999. Our results suggest that earnings multiples became more similar over this time period for the jurisdictions we analyze, although systematic differences remain. Economic determinants of earnings multiples (e.g., growth rates, interest rates, and returns) do not exhibit similar convergence and do not appear to explain the changes. The convergence is robust to controls for cash flow multiples and is apparent in the valuation of accruals. Accrual/cash flow correlations have also become more similar and generally less negative, suggesting a reduction in earnings smoothing. Overall, our evidence suggests convergence in accounting practice.

Keywords: *international accounting; international valuation; accounting harmonization; accounting multiples.*

Data Availability: *The data used in this study are publicly available from the sources indicated in the text.*

I. INTRODUCTION

We document convergence in earnings multiples for a sample of firms from Australia, Canada, France, Germany, Japan, the United Kingdom, and the United States over the years 1987–1999. Convergence persists after including controls for earnings, sales and GDP growth rates, interest rates, and returns, and is robust to inclusion of cash flow multiples. We obtain similar results focusing on accruals multiples, suggesting that the pricing of accruals drives the empirical results. We document that cash flow/accrual correlations have become more similar for our sample, and have become less negative for all countries other than the U.S. and Canada. Finally, we provide evidence that

---

We thank Katherine Schipper (editor), David A. Guenther (discussant), Allison Evans, Maria Nondorf, two anonymous reviewers, and workshop participants at *The Accounting Review*'s Conference on Quality of Earnings, Emory University, University of Maryland, University of Michigan Spring Training, and The University of North Carolina at Chapel for helpful comments on earlier versions.

book value multiples have converged, although the evidence is weaker than for earnings. Taken together, our results suggest a reduction in cross-country accounting differences over time.

There are several reasons to expect cross-country convergence in accounting. Most directly, there have been significant efforts to reduce accounting differences and develop consistent regulation worldwide through organizations like the International Accounting Standards Committee (IASC), International Auditing Practices Committee, International Organization of Securities Commissions, European Union, and Association of Southeast Asian Nations. Countries like the U.S., Canada, and the U.K. have worked together and with the IASC to develop joint standards, and auditing firms have been active in forming international alliances to encourage consistent practice across countries.

Further, cross-border competition for capital has created incentives to improve accounting quality and comparability. Anecdotal evidence suggests that some large companies have voluntarily switched accounting principles to be more consistent with international norms, and companies have retained international auditing firms to provide greater confidence in their reports.[1] As a result, accounting practice appears to have become more consistent internationally, leading some to conclude that a continuation of the process currently in place with incentives for firms to voluntarily move toward international norms can result in high-quality, comparable accounting practice going forward.[2] For example, the NYSE and others have argued that relaxed cross-listing standards and increased reliance on IAS may be appropriate if cross-border accounting differences are no longer substantial (especially for issuers that would consider listing in the U.S.) and are getting smaller (Cochrane 1994; Cochrane et al. 1996).

Others view progress as more apparent than real, believe that substantial differences remain, and express concern about the perception that differences in accounting have been substantially reduced. The FASB, for example, tallied 255 differences between IASC standards and U.S. GAAP, and concluded that many important differences still existed internationally (FASB 1996). Similarly, Choi et al. (1999), and others, suggest that important accounting differences remain across countries, although the differences identified are generally in alternatives permitted rather than explicit differences in practice.

In the end, the extent of convergence of financial reporting internationally is an empirical question. Our goal is to provide descriptive evidence on whether accounting differences, as reflected in differences in earnings multiples, appear to have decreased over time and, if so, what factors explain the change. Our primary focus is on earnings multiples because earnings is a key summary performance measure in all of our sample countries and because earnings (as a summary measure) should be strongly affected by changes in accounting practice.

While we view our analysis as descriptive, we are guided by two constructs of interest to regulators: "comparability" and "quality." Both constructs are subject to differences in interpretation and are difficult to operationalize, so we do not attempt to measure them directly. Like much of the literature, we take an indirect approach, based on the valuation characteristics of the resulting net income, to infer changes in underlying accounting practice. What we have in mind for comparability is the idea expressed by Jenkins (1999), that "what we all want and need is a set of high-quality standards that assures that similar transactions are accounted for similarly and that unlike transactions are not." We base our assessment of comparability on the multiples applied to accounting data by investors, and argue that increased comparability will be reflected in more similar multiples, all else equal. Our accounting quality construct is a measure of smoothing based on the correlation of cash flows and accruals; we argue that less smoothing implies increased earnings quality.[3]

---

[1] David Cairns, former Secretary-General of the IASC, states, for example, "many European companies have recognized the need to make their financial reports more relevant to, and understandable by, international capital markets. They have responded to this need by going well beyond the limited requirements of the national accounting laws and standards of the time" (Cairns 1997). Choi et al. (1999) make a similar case.

[2] Cairns (1997) notes, for example, "when we look at the way countries or companies account for particular transactions or events, it is increasingly difficult to distinguish in a systematic way so-called Anglo-American accounting from Continental European accounting or American accounting from, say, German accounting." Similarly, Choi et al. (1999) indicate, "harmonization debates aside, all dimensions of accounting *are* being harmonized worldwide" (emphasis in the original).

Our indirect approach takes into account the effects of actual accounting practices, from the perspective of users of accounting information, by basing analyses on investor valuations of reported earnings. We do not attempt to examine accounting practices directly. The extent to which companies exploit allowable alternatives and the magnitude of the resulting effect on income, while potentially large, is generally not observable, and studies like Ball, Robin, and Wu (2000) and Pownall and Schipper (1999) note that accounting changes often result from incentives rather than regulation. The strength of auditing and enforcement also varies across countries, potentially resulting in differences in the application of accounting standards. Further, it is difficult to judge accounting practice based on accounting policy footnotes since disclosure is typically vague, especially outside the U.S., and covers, at most, major accounting choices (Frost and Ramin 1997).

However, our indirect approach creates the potential for omitted correlated variable bias. To counter that possibility, we include tests based on recent approaches to measure earnings management; we also identify where we expect effects to be most pronounced and in what direction, and attempt to incorporate likely correlated variables into our empirical analysis. Further, systematic differences across countries should be mitigated by the fact that we conduct our comparisons over time.

We interpret our findings as most consistent with the notion that the convergence in earnings multiples reflects, at least in part, convergence in accounting practices over time, but that systematic accounting differences remain. Further, our evidence on the time-series properties of earnings suggests that smoothing has been mitigated, and has become more similar across our sample countries. While we cannot dismiss the possibility of omitted correlated variables, these results are robust to a variety of controls.

In the next section, we describe the data. Then, we present the analysis of changes in earnings multiples, followed by our earnings smoothing tests and supplementary tests based on the multiple applied to book value. We conclude with a summary and discussion of caveats.

II. DATA

Our data are from the 1997 and 2000 Global Vantage disks, which is the most comprehensive dataset with the longest time-series that we could find.[4] Our analysis excludes firms in developing economies because they generally lack a sufficient time-series to permit inference, but includes firms in countries that have historically been at the extremes in terms of reporting philosophy. Further, our sample countries all have developed, diversified economies, so firms should face fairly similar economic environments. We exclude financial firms, because they face unique reporting issues and there are not enough to permit a separate analysis.

---

[3] While these goals have some intuitive appeal, we do not claim that they are necessarily optimal or that our constructs capture them perfectly. Further, our approach focuses on measurement issues and their effects on bottom-line net income. Supplemental disclosure and other financial statement issues are also important, but are even more difficult to assess and have not (at least to date) been the subject of as much concern as income measurement issues.

[4] Some studies (e.g., Alford et al. 1993) use data for a few years earlier, but licensing agreements and transfers of ownership between data providers limit our ability to access those data. We compared Global Vantage, Datastream, and Worldscope, but data availability (especially for early years) is best with Global Vantage. We avoided mixing data providers to increase the likelihood that the data are prepared on a consistent basis.

Table 1 provides descriptive statistics for our sample firms. In the average year, our sample includes 130 Australian firms, 282 Canadian firms, 168 German firms, 172 French firms, 601 U.K. firms, 902 Japanese firms, and 1,940 U.S. firms. Not surprisingly, the larger economies with a longer history of equity ownership tend to have more firms with available data. Because the number of firms varies across subperiods, we replicated tests limiting the

TABLE 1
Descriptive Statistics for Sample Firms

	Number	*Median*	*Mean*	*1st Qu.*	*3rd Qu.*
Panel A: Group 1 (Years 1987–1992)					
Earnings/Price					
Australia	657	.074	.096	.047	.106
Canada	1,601	.062	.071	.035	.091
Germany	682	.038	.042	.024	.057
France	580	.064	.080	.046	.096
United Kingdom	2,991	.076	.089	.052	.109
Japan	4,190	.017	.020	.010	.027
United States	11,713	.064	.073	.041	.092
Book/Market					
Australia	652	0.65	0.87	0.44	1.03
Canada	1,563	0.72	0.88	0.47	1.06
Germany	673	0.45	0.52	0.29	0.67
France	574	0.50	0.90	0.34	0.81
United Kingdom	2,953	0.56	0.75	0.35	0.95
Japan	4,149	0.27	0.30	0.17	0.40
United States	11,219	0.56	0.68	0.34	0.85
Total Assets ($ Millions)					
Australia	642	288	854	103	919
Canada	1,563	295	929	105	1,028
Germany	663	632	2,783	200	2,333
France	565	1,276	3,083	514	3,474
United Kingdom	2,925	217	925	82	755
Japan	4,105	1,008	3,078	412	2,742
United States	11,472	252	1,262	93	971
Return on Equity					
Australia	640	11.7	11.0	6.4	16.9
Canada	1,560	9.7	8.6	3.7	14.7
Germany	661	8.5	8.7	5.0	12.5
France	564	13.7	13.6	8.7	18.4
United Kingdom	2,917	15.2	16.3	7.4	22.8
Japan	4,083	6.8	7.3	4.6	9.4
United States	11,290	12.5	11.5	6.3	18.2

(*Continued on next page*)

TABLE 1 (Continued)

	Number	*Median*	*Mean*	*1st Qu.*	*3rd Qu.*
Panel B: Group 2 (Years 1994–1999)					
Earnings/Price					
Australia	897	.063	.072	.041	.087
Canada	1,777	.052	.066	.030	.082
Germany	1,339	.045	.068	.027	.076
France	1,485	.055	.066	.035	.085
United Kingdom	4,224	.063	.078	.042	.095
Japan	6,633	.031	.038	.017	.050
United States	11,564	.053	.061	.033	.077
Book/Market					
Australia	881	0.59	0.67	0.40	0.84
Canada	1,726	0.60	0.73	0.35	0.91
Germany	1,316	0.47	0.57	0.29	0.69
France	1,467	0.55	0.70	0.32	0.91
United Kingdom	4,065	0.44	0.59	0.24	0.78
Japan	6,560	0.76	0.91	0.51	1.16
United States	10,983	0.45	0.56	0.26	0.71
Total Assets ($ Millions)					
Australia	874	263	809	88	758
Canada	1,737	341	1,084	141	1,084
Germany	1,307	423	2,785	139	1,519
France	1,449	370	2,554	122	1,727
United Kingdom	4,132	194	1,037	71	726
Japan	6,497	769	2,553	355	2,091
United States	11,344	547	2,091	185	1,765
Return on Equity					
Australia	873	10.5	10.0	6.3	15.5
Canada	1,723	10.2	9.6	4.6	15.5
Germany	1,308	10.7	10.8	5.6	16.2
France	1,448	11.7	11.4	6.7	16.7
United Kingdom	4,126	15.1	15.7	7.8	24.4
Japan	6,504	3.6	3.1	1.5	6.0
United States	11,107	12.8	12.2	6.9	19.1

Data are from Global Vantage covering 1987–1999. Earnings/Price is earnings before extraordinary items divided by year-end market value and Book/Market is shareholders' equity divided by year-end market value, both with negative values excluded and extreme values truncated at 1 percent. Return on equity is earnings before extraordinary items divided by shareholders' equity.

sample to firms with data in both periods. While sample sizes are smaller, especially for some of the smaller countries, results are consistent. Japanese, German, and French firms tend to be larger, while U.K., U.S., Canadian, and Australian firms tend to be smaller, but in each case there is a substantial overlap in the distribution of size across countries. Results that follow are robust to controls for size.[5]

---

[5] Observations in each country are generally spread across a range of industries and, because the comparison is over time, cross-country differences in industry concentration should not affect results. Including industry controls does not affect our conclusions.

III. CHANGES IN MULTIPLES OVER TIME

Table 1 presents earnings/price ratios by country. Following the conventions in Global Vantage, earnings/price is defined as earnings before extraordinary items divided by market value at year-end. We begin with a potential sample of 62,378 firm years and use the same deletion rules as Joos and Lang (1994). Because earnings/price ratios are difficult to interpret for loss firms, we delete observations with negative earnings (18 percent of potential observations in the first period and 20 percent in the second period) and, to mitigate the effects of extreme observations, we delete the top 1 percent of earnings/price ratios, leaving 50,333 observations.[6]

The overall pattern in earnings/price ratios is generally consistent with that reported in other studies, beginning with Speidell and Bavishi (1992). Earnings/price ratios are lowest in Japan and Germany, consistent with their legal origins as code-law countries with close links between financial reporting and tax reporting, and a historical creditor focus, resulting in more conservative income measurement.[7] Consistent with the differences in multiples reflecting the effects of traditional accounting conservatism, median return on equity for Japan and Germany is the lowest among our sample countries during the 1987–1992 subperiod. Similarly, GDP growth rates (not tabulated) during the sample period are no higher for Japan and Germany than for other sample countries, suggesting that high growth does not explain the higher multiples. Similarly, earnings/price ratios and return on equity are generally higher in the common law countries (Australia, Canada, U.K., and U.S.), consistent with those countries following more equity-focused, less tax-based accounting and, consequently, exhibiting less conservative accounting practices. The ordering across countries is consistent with the Barth and Clinch (1996) finding that Form 20-F net income reconciling items are on average negative for Australia, Canada, and U.K., and largest for Australia.

To provide initial descriptive evidence on changes in earnings multiples over time, we estimate a regression after splitting the sample period into two subperiods, comprising the first six years and the last six years.[8] We regress mean-adjusted earnings/price ratios on 14 indicator variables for each country/subperiod grouping (i.e., seven countries multiplied by two subperiods). We adjust each earnings/price ratio by subtracting the average (over firms, years, and countries) earnings/price ratio for that subperiod (.067 for the first subperiod and .064 for the second subperiod) so that the coefficient estimates in the regression represent deviations for a given country from the mean.[9]

---

[6] Results are very similar if loss observations are included. In addition, results are robust to analysis based on medians and rank regression, suggesting that remaining large observations do not drive results.

[7] France is more difficult to categorize. Based on simulations of European accounting practices, Simmonds and Azieres (1989) find that Germany is consistently more conservative than the U.K., but France generally lies between Germany and the U.K., and in some cases is less conservative than the U.K. (see also Alexander and Archer 1991). Ball, Kothari, and Robin (2000) also find that France lies between Germany and the U.K. in terms of earnings valuation.

[8] We use groupings to increase the power of the tests. As we discuss later, results are robust to alternate groupings and are consistent on a year-by-year basis, suggesting that a subset of years does not drive the results. We use a regression specification for consistency with later analysis when we include controls for other potential factors. Results are similar for a standard t-test comparison of means.

[9] This approach implicitly assumes that the relevant benchmark for assessing convergence is whether countries have become more similar. An alternate approach would be to assess whether accounting worldwide has moved toward a benchmark set of practices, like that represented by U.S. GAAP. Redefining convergence as movement toward U.S. multiples yields very similar results because the U.S. is close to the average and stable over the sample period.

Table 2 contains the regression results. In general, a comparison between periods suggests convergence in the sense that dispersion is reduced. In addition, convergence is apparent for all seven sample countries, with each moving toward the mean in the second period. For example, the mean earnings/price ratio for Australia, which was .0284 greater than the mean in the first period, is .0078 greater than the mean in the second period. Earnings/price multiples in Canada, France, U.K., and U.S. also move from above the mean toward the mean. Finally, multiples in Germany and Japan, which were below the mean in the first period, move upward toward the mean in the second period. An F-test of the difference in coefficient estimates between the first and second periods indicates the shift is significant at the .001 level for each country but Canada. However, multiples in most countries remain significantly different from the mean in the second period, indicating that whatever forces caused divergence in the first period have been reduced but not eliminated.[10]

In general, the results suggest a narrowing of differences in earnings multiples over the sample period for all countries, with results strongest in the traditional code-law countries. Subject to concerns about other potential economic determinants, this result is consistent with the assertion by many commentators that, worldwide, accounting differences have been narrowed and that much of the recent progress in harmonizing accounting has occurred in

---

[10] An issue with the statistical inference is potential autocorrelation of residuals. Residual autocorrelation is approximately 0.3 and applying a Cochrane-Orcutt adjustment does not affect the inference.

TABLE 2

Regression Results for Earnings to Price Ratios on Country/Subperiod Indicators

Country	*Subperiod*	*Coefficient*	*t-statistic*	*Significance*	*Change Significance*
Australia	1st	.0284	14.13	.001	
	2nd	.0078	4.55	.001	.001
Canada	1st	.0034	2.65	.008	
	2nd	.0020	1.61	.107	.417
Germany	1st	−.0248	−12.55	.001	
	2nd	.0040	2.84	.005	.001
France	1st	.0129	6.03	.001	
	2nd	.0020	1.51	.131	.001
U.K.	1st	.0219	23.20	.001	
	2nd	.0134	16.83	.001	.001
Japan	1st	−.0476	−59.77	.001	
	2nd	−.0264	−41.69	.001	.001
U.S.	1st	.0058	12.09	.001	
	2nd	−.0028	−5.84	.001	.001

Results are from the regression of mean-adjusted earnings/price ratios on 14 country/subperiod indicator variables, with no intercept. Earnings are before extraordinary items and price is as of year-end. The subperiods are 1987–1992 and 1994–1999, and there are 50,333 observations. Earnings/price ratios are adjusted by subtracting the subperiod mean of .067 in the first subperiod and .064 in the second. The change significance is from an F-statistic for the change in coefficient estimates between periods.

code-law countries, with standard setters adopting more international norms and companies going beyond national requirements to attract capital.[11] Further, the results are consistent with the concern expressed by some (e.g., Biener 1994) that accounting worldwide has moved toward a common law, equity-focused perspective to the detriment of other stakeholders. Similarly, the evidence on earnings multiples for Australia is consistent with its historic reputation for aggressive accounting and its more recent movement toward IAS (*Economist* 1999).

An alternative test for the narrowing of differences is to compare the cross-country coefficient of variation of average earnings/price ratios in each year (i.e., seven observations to each variance).[12] Comparing the first six years in the sample period to the last six provides striking evidence of a reduction in dispersion (results not tabulated). Each yearly coefficient of variation in the first period is larger than the corresponding coefficient in the second period. In fact, the smallest of the six first-period annual coefficients of variation is larger than the largest of the six second-period coefficients of variation, suggesting that the relation is consistent and is not driven by a subset of years. Applying a rank sum test comparing the first and second periods indicates that the reduction in average variance is significant at the .01 level. The rank correlation between coefficients of variation and time is -0.87, significant at the .001 level with only 13 observations.

A second alternative approach is to compare the proportion of the variation in multiples that can be explained by country differences across the two periods. When we estimate a regression of earnings/price ratios on the country indicator variables (omitting one country to be captured in the intercept), the R^2 measures the extent to which country-specific differences explain multiples. In the first period, the R^2 is .17 vs. .07 in the second period, indicating that country-specific factors are more important in the first period (results not tabulated). The difference in R^2s is significant at the .01 level based on a Cramer Z-statistic.

Because earnings multiples are a function of economic factors like expected growth and risk, the results based on these tests may simply reflect convergence in other economic factors. We take several approaches to address this possibility. First, we replicate the preceding analysis for cash flow multiples. To the extent that economic factors and not accounting practices are important, they should also affect the multiples applied to cash flows. Because statements of cash flows are not available for most of our firms during the sample period, we estimate cash from operations using net income adjusted for changes in balance sheet accounts.[13] Results for cash-flows-to-price ratios, reported in Table 3, show no consistent evidence of either convergence or divergence between the two periods. Cash flow multiples for Canada, Germany, and U.S. move away from the average, and Australia and France show no change. The fact that cash flow multiples change between periods is not surprising, since factors like expected growth and risk may change as well. Given the results in Table 3, we include controls for cash flow multiples in the analysis that follows.

---

[11] See, for example, Choi et al. (1999) for a discussion. For a discussion of progress toward consistent standards for specific countries during the sample period see Curran (1996) on Australia; *CA Magazine* (1996) on Canada; Doupnik (1992) and Pape (1997) on Germany; *Accountancy* (1997a) on France; Shibata (1998) on Japan; and *Accountancy* (1997b) on the U.K.

[12] The coefficient of variation is the standard deviation of the seven-country mean earnings/price ratios divided by the grand mean of the seven-country means to control for movements in average earnings/price ratios over time. Results are not sensitive to dividing by the mean.

[13] The calculation is net income before extraordinary items plus depreciation minus the change in noncash assets (other than fixed assets) plus the change in liabilities (other than long-term debt). The decreased number of observations reflects missing data. To mitigate the effect of extreme observations, accruals in excess of market value in absolute value were deleted (these are generally large negative accruals consistent with "big baths").

TABLE 3
Regression Results for Cash Flows to Price Ratios on Country/Subperiod Indicators

Country	*Subperiod*	*Coefficient*	*t-statistic*	*Significance*	*Change Significance*
Australia	1st	−.0213	−2.81	.005	
	2nd	−.0248	−3.99	.001	.726
Canada	1st	.0061	0.99	.325	
	2nd	−.0156	−3.13	.002	.006
Germany	1st	.0216	2.95	.003	
	2nd	.0397	7.95	.001	.042
France	1st	.0576	6.67	.001	
	2nd	.0506	10.05	.001	.483
U.K.	1st	.0169	4.57	.001	
	2nd	−.0151	−5.19	.001	.001
Japan	1st	−.0783	−22.67	.001	
	2nd	−.0130	−4.53	.001	.001
U.S.	1st	−.0025	−1.31	.190	
	2nd	−.0218	−12.21	.001	.001

Results are from the regression of mean-adjusted cash flows/price ratios on 14 country/subperiod indicator variables, with no intercept. The subperiods are 1987–1992 and 1994–1999, and there are 50,487 observations. Cash flows from operations/price ratios are adjusted by subtracting the subperiod mean of 0.115 in the first subperiod and 0.113 in the second. The change significance is from an F-statistic for the change in coefficient estimates between periods.

Our second approach to controlling for economic effects on earnings multiples is to explicitly control for several factors identified by previous research that might affect earnings multiples. French and Poterba (1991) and Guenther and Young (2000) suggest that differences in earnings multiples might reflect differences in GDP growth or real interest rates. To examine that possibility, we collected data on GDP growth rates from the OECD Historical Statistics and on real interest rates from the IMF World Economic Outlook. Similarly, we gathered earnings and sales growth rates for each firm over the sample period to control for the possibility that convergence in multiples might reflect changes in firm-specific growth.[14] Finally, we collected data on prior and contemporaneous returns to investigate if the change in multiples reflects the effects of past or future returns due to some combination of misvaluation and deviations between forecasted and actual growth.

We first examined the control variables for evidence of patterns over time. To the extent that the pattern in earnings multiples is being driven by convergence in economic factors captured by the control variables, there should be evidence of convergence in GDP, earnings and sales growth rates, interest rates, and returns. Plots of each control variable (not tabulated) in each subperiod provide no evidence of convergence in any of the measures. The dispersion of each of the firm-specific variables increases between periods, suggesting that earnings multiple results are not driven by firms becoming generally more similar over the sample period.

---

[14] We were unable to obtain growth forecasts for most of our sample firms, especially in the first period, and use realized rates instead. Including returns in the regression should partially capture the effects of differences between forecasted and realized growth. In addition, we replicated the analysis using year-ahead and two-year-ahead realized earnings growth. While that approach limits our sample period by eliminating the later years, results are very similar.

More to the point, including the controls in the regressions does not affect conclusions. Table 4 presents results for mean-adjusted earnings/price ratio regressions with controls included. The number of observations decreases because of missing data on returns or the change in earnings. However, the pattern of results is similar to that in Table 2, with all countries except Canada moving significantly toward the mean. Statistical tests yield similar conclusions as before; based on a rank sum test, the cross-country variance of earnings/price ratios decreases significantly (at the .01 level) between periods after controlling for these potential economic determinants.

TABLE 4
Regression Results for Earnings/Price on Country/Subperiod Indicators with Controls for GDP Growth, Real Interest Rates, Earnings Growth, and Returns

Country	*Subperiod*	*Coefficient*	*t-statistic*	*Significance*	*Change Significance*
Australia	1st	.0278	8.24	.001	
	2nd	.0102	3.63	.001	.001
Canada	1st	.0019	0.72	.474	
	2nd	.0042	1.62	.104	.193
Germany	1st	−.0213	−7.94	.001	
	2nd	−.0001	−0.03	.978	.001
France	1st	.0116	3.43	.001	
	2nd	.0030	1.25	.212	.003
U.K.	1st	.0195	12.33	.001	
	2nd	.0141	6.47	.001	.001
Japan	1st	−.0446	−26.70	.001	
	2nd	−.0319	−25.93	.001	.001
U.S.	1st	.0051	3.08	.002	
	2nd	.0004	0.24	.812	.001
GDP Growth		−.0016	−9.00		
Int. Rate		−.0002	−0.58		
EPS Growth		−.0028	−28.03		
Return		.0000	0.63		

Results are from the regression of mean-adjusted earnings/price ratios on 14 country/subperiod indicator variables and controls for GDP growth, real interest rates, earnings growth, and returns, with no intercept. The subperiods are 1987–1992 and 1994–1999, and there are 42,300 total observations. Earnings/price ratios are adjusted by subtracting the subperiod mean of .074 in the first period and .069 in the second. The change significance is from an F-statistic for the change in coefficient estimates between periods.

We again compare the explanatory power of country indicators using a two-stage approach. First, we regress earnings/price ratios on the economic controls to capture the effects of these economic factors by period. Second, we regress the residuals from the first stage on the country indicator variables to compare the extent to which country-specific factors explain the residual variation in multiples in the two periods. As before, the difference in R^2s (not tabulated) is significant at the .01 level, indicating that, even controlling for several economic variables, the importance of country-specific variables in explaining the dispersion in earnings multiples has decreased over time. While our control variables are likely measured with error, they should be correlated with the "true" economic determinants, yet our results are at least as strong with the controls as without.

Another possibility is that convergence in earnings multiples reflects risk shifts or changes in market segmentation. While we know of no empirical evidence of risk shifts during the sample period, this possibility is difficult to dismiss because the nature of priced risk is not well understood, especially in international contexts. We do not have sufficient returns data to estimate a capital asset pricing model for our early period and it is not clear, even if we did, which pricing model is appropriate for our purposes (a one-factor model using each country's index assumes away any effect since the average beta is constrained to be 1). Following Dhaliwal et al. (1999), we use leverage as our proxy for risk. In general, there is no evidence that leverage became more similar during the sample period and including leverage as a control does not affect our conclusions (results not tabulated). Further, for a subset of firms with monthly returns data on Global Vantage, we estimate betas relative to local and world average returns. While the resulting betas are very noisy due to the small number of observations, there is no evidence of convergence in those estimates and our results are robust to their inclusion.

A related possibility is that earnings multiples are converging simply because the underlying economics of our sample firms are becoming more similar. If that is true, then within-country dispersion should also decrease over time. To examine that possibility, we compare the annual within-country coefficient of variation over time for the sample firms. Because data are available for a much longer time for the U.S. firms and trends may be more apparent over longer periods, we also compare the coefficients of variation over the period from the 1950s to 1990s for U.S. firms. Results from this analysis (not tabulated) provide no evidence of a general reduction in variation over time. Of the seven sample countries, four experienced increases in dispersion over the sample period. The U.S. analysis suggests an increase in dispersion over time, primarily in the late 1970s, although the pattern is basically flat during our sample period, which covers 1987–1999.

The preceding analysis includes variables that previous research has shown are likely to influence our results, but it is possible that potentially important unidentified variables are excluded. To the extent that such unidentified factors affect the earnings multiples, they should also be reflected in the multiples applied to cash flows. As noted above, there is some evidence that cash flow multiples change between periods, so controlling for the cash flow multiple may mitigate the effects of omitted economic factors on earnings multiples.[15] We include the ratio of cash flow to price, interacted with indicator variables for each country/subperiod, to control for changes in cash flow multiples between periods.[16]

---

[15] Another possibility is that changes in earnings multiples are related to changes in the market multiple applied to the book value of equity. Results are robust to including controls for changes in the multiple on book value.

[16] Specifically, we include 14 control variables, one for each country/subperiod, which take on a value of the cash flow/price ratio if the observation is for that country in that subperiod, and 0 otherwise. This allows the coefficient to vary between countries and between subperiods.

Changes in expected profits or risk should be at least partially captured in the cash flow multiple, and controlled by the inclusion of cash flows in the regression.

Results including both the economic controls and cash flow multiples are reported in Table 5. The results are very consistent with those reported previously, indicating that controlling for cash flow multiples does not change the conclusion that earning multiples have converged over time. With the exception of France and Canada (which were not significantly different from the mean in the first period), the change in coefficient estimates is significant and the rank sum tests (not tabulated) indicate convergence. While there is substantial evidence of convergence, however, the coefficient estimates in the second period generally remain significantly different from the mean, indicating that convergence is incomplete. Again, we apply a two-stage approach in which we first condition on the economic and cash flow controls and then assess the ability of the country indicator variables to explain remaining variation (results not tabulated). As before, there is strong evidence

TABLE 5

Regression Results for Earnings/Price on Country/Subperiod Indicators with Controls for GDP Growth, Real Interest Rates, Earnings Growth, Returns, and Cash Flow/Price

Country	*Subperiod*	*Coefficient*	*t-statistic*	*Significance*	*Change Significance*
Australia	1st	.0280	7.61	.001	
	2nd	.0104	3.37	.001	.001
Canada	1st	.0015	0.48	.629	
	2nd	.0024	0.79	.427	.707
Germany	1st	−.0223	−7.64	.001	
	2nd	−.0038	−1.58	.113	.001
France	1st	.0047	1.16	.245	
	2nd	.0035	1.28	.201	.730
U.K.	1st	.0228	13.08	.001	
	2nd	.0151	6.06	.001	.001
Japan	1st	−.0415	−19.71	.001	
	2nd	−.0284	−19.99	.001	.001
U.S.	1st	.0068	3.64	.001	
	2nd	.0009	0.41	.683	.001
GDP Growth		−.0014	−7.17		
Int. Rate		.0000	0.08		
EPS Growth		−.0031	−28.00		
Return		.0000	2.11		

Results are from the regression of mean-adjusted earnings/price ratios on 14 country/subperiod indicator variables and controls for GDP growth, real interest rates, earnings growth, returns, and 14 controls for country and period cash flows from operations/price (not tabulated), with no intercept. The subperiods are 1987–1992 and 1994–1999, and there are 33,710 total observations. Earnings/price ratios are adjusted by subtracting the subperiod mean of .068 in the first subperiod and .066 in the second. The change significance is from an F-statistic for the change in coefficient estimates between periods.

(significant at the .01 level) that the ability of country-specific factors to explain remaining variation is lower in the second period than in the first.

The results in Tables 3 and 5 indicate that the pattern in earnings multiples is robust to controls for changes in cash flow multiples, suggesting that the pricing of accruals drives the results. We test this implication by replicating the analysis of earnings multiples using accruals deflated by price as the dependent variable, including controls for cash flows and other economic factors. Table 6 reports coefficients on accruals/price regressions analogous to those in Table 5, but with earnings replaced by accruals.

The pattern of convergence in accruals is similar to that for earnings, suggesting that the pricing of accruals is driving the changing multiples on earnings. For all countries except Canada and France, the multiple moves closer to the mean in the second period and, in every case in which the multiple was significantly different from the mean in the first period, the change is statistically significant at the .01 level or better. Further, for the cases in which the first period coefficient is not reliably different from the mean, there is evidence of a significant change between periods only for the U.S., suggesting that changes occur only in cases in which they would be predicted. Finally, in most cases the multiple retains its sign relative to the mean in the second period, indicating that accounting influences on earnings multiples in the first period were mitigated, but not eliminated. Applying the rank sum test described earlier, the reduction in dispersion is statistically significant at

TABLE 6

Regression Results for Accruals/Price on Country/Subperiod Indicators with Controls for GDP Growth, Real Interest Rates, Earnings Growth, Returns, and Cash Flow/Price

Country	*Subperiod*	*Coefficient*	*t-statistic*	*Significance*	*Change Significance*
Australia	1st	.0273	4.26	.001	
	2nd	.0011	0.21	.832	.001
Canada	1st	.0033	0.59	.553	
	2nd	−.0004	−0.07	.946	.365
Germany	1st	−.0189	−3.74	.001	
	2nd	−.0020	−0.47	.641	.001
France	1st	−.0006	−0.09	.925	
	2nd	.0051	1.09	.275	.341
U.K.	1st	.0296	9.76	.001	
	2nd	.0169	3.91	.001	.001
Japan	1st	−.0447	−12.23	.001	
	2nd	−.0212	−8.59	.001	.001
U.S.	1st	.0040	1.23	.217	
	2nd	.0004	0.10	.923	.008

Results are from the regression of mean-adjusted total accruals/price ratios on 14 country/subperiod indicator variables and controls for GDP growth, real interest rates, earnings growth, returns, and 14 controls for country and period cash flows from operations/price, with no intercept. The subperiods are 1987–1992 and 1994–1999, and there are 33,914 total observations. Accruals/price ratios are adjusted by subtracting the subperiod mean of −.049 in both the first and second subperiods. The change significance is from an F-statistic for the change in coefficient estimates between periods.

the .01 level. Comparing R^2s, the ability of cross-country differences to explain variation in the residual after controlling for the other factors is significantly (at the .01 level) lower in the second period than the first.

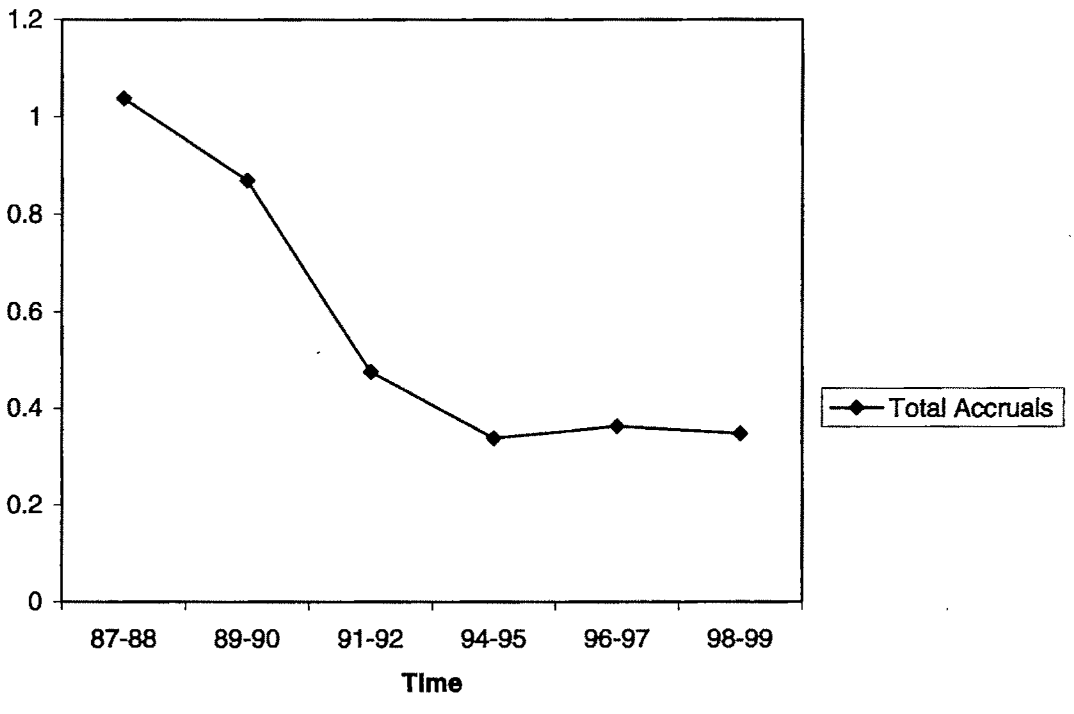
To provide a better sense for the pattern in reduction in dispersion over time, Figure 1 presents the coefficient of variation of accruals multiples after controlling for cash flows and economic variables (analogous to Table 6) over the sample period. Data are grouped into two-year subperiods to increase the number of observations for each country in each subperiod. The pattern in Figure 1 suggests a fairly consistent reduction over the first two-thirds of the total period. While there is less evidence of a reduction in the last third of the sample period, evidence of convergence is not limited to a subset of periods or to one or two countries.[17]

While we are generally agnostic as to the cause of the observed convergence in earnings and accruals multiples, we provide additional analyses of three possible causes: changes in consolidation policy over time, an increase in the proportion of our sample firms listed on

---

[17] The pattern is similar for the standard deviation plotted over time, but deflating by the mean ensures that the pattern is not driven by changes in average multiples over time. Results are consistent for a nonparametric approach using the interquartile range divided by the median. A similar analysis based on earnings/price ratios indicates reductions in dispersion in the later portion of the sample period (although less pronounced) as well as in the earlier period.

FIGURE 1
Coefficient of Variation of Accruals Multiples



Coefficients of variation of the coefficients are from the regression of mean-adjusted total accruals/price ratios on 42 country/subperiod indicator variables and controls for GDP growth, real interest rates, earnings growth, returns, and cash flows from operations/price. The data are grouped into two-year subperiods.

U.S. markets, or an increase in the number of our sample firms reporting under nondomestic standards. Based on the Global Vantage code for consolidation policy, the proportion of firms filing unconsolidated financial statements decreases from 9.0 percent of sample firms in the first period to 4.5 in the second period. However, results are consistent excluding unconsolidated firms.

The second possible explanation is an increase in the number of sample firms trading on U.S. markets. Although Global Vantage reports data under local GAAP, research like Lang et al. (2002) suggests that firms alter their local-GAAP accounting choices around the time of a cross-listing. To assess the extent to which our sample firms are cross-listed on U.S. markets, we compare our sample with a list of cross-listed firms from Compustat. The proportion of sample firms cross-listed on U.S. markets remains at about 8.0 percent between the two subperiods, and excluding those cross-listed firms does not affect our conclusions.

Finally, it is possible that more sample firms are choosing to report under standards other than local GAAP. Based on Global Vantage reporting codes, the proportion of sample firms reporting under IAS or U.S. GAAP has increased over time, but it averages only 0.8 percent and 2.1 percent in the first and second sample periods, respectively, and the use of IAS or U.S. GAAP is concentrated in Germany and France. Conclusions are not sensitive to excluding those observations. Taken together, these results suggest that changes in consolidation policy, cross-listing, and reporting under nondomestic standards do not explain our empirical results.

IV. CHANGES IN EARNINGS SMOOTHING

Because our analysis so far relies on share price to detect accounting differences, it is possible that our results may be driven by factors like market inefficiency, market segmentation, or inadequate controls for risk or expected growth. In this section, we assess convergence by looking to the time-series properties of earnings, without reference to share price.

A frequently cited worldwide concern is the ability of firms to smooth earnings by setting up reserves in good times to be drawn down later. While reserves are an issue everywhere, the rules regulating them are probably most stringent in the U.S., where they have been an ongoing priority for the SEC. Outside the U.S., especially in code-law countries, there has been less emphasis on regulating reserves, perhaps because understating net assets is viewed as less of a concern than overstating them. Foreign registrants frequently reverse reserves when reconciling their financial statements to U.S. GAAP and firms in Germany and Japan typically report lower earnings variability, consistent with smoothing (Leuz et al. 2001; Ball, Kothari, and Robin 2000).

One approach to assessing the effects of income-smoothing behavior is based on the correlation between accruals and cash flows. Authors like Myers and Skinner (2001), Beatty et al. (2002), and Leuz et al. (2001) have argued that, all else equal, earnings smoothing should induce a more negative correlation between cash flows and accruals. For example, Leuz et al. (2001) find that measures based on this correlation approach are highly correlated with other measures of earnings management, and differ predictably across countries based on the importance of equity markets, ownership concentration, shareholder protection, disclosure, and enforcement.[18]

---

[18] In an earlier draft, we also included an analysis based on the proportion of small positive earnings as in Burgstahler and Dichev (1997). That approach is more difficult to interpret since it assumes that firms in all countries are managing toward the same objective, but conclusions were generally consistent with those presented here.

While there may be intrinsic country-specific economic factors that cause the relation between cash flows and accruals to vary, we know of no reason to expect the underlying economics to drive the cash flow/accruals correlation in any systematic direction, especially within countries over time. Further, research like Leuz et al. (2001) finds cross-country results using this measure that are consistent with predictions.

Table 7 presents results for the correlations over time based on accruals and cash flows as calculated in the previous section, deflated by total assets to mitigate the effect of size.[19] We use rank correlations to reduce potential effects of extreme observations, but results are consistent for Pearson correlations. In general, the cross-country ordering of correlations is consistent with results in prior research. Germany, Japan, and France have the most negative correlations, suggesting that earnings smoothing is more pervasive in those environments. Further, the magnitudes are striking, with correlations of about −90 percent in each case. Correlations are less negative, but still quite large, in the U.K. and Australia. The U.S. exhibits the smallest correlation at −56 percent with Canada close behind at −66 percent, suggesting that earnings smoothing in the U.S. is less pronounced than in the other leading market economies.[20]

Most countries experience a significant reduction in the negative correlation between cash flows and accruals between the first and second sample periods. Following Beatty et

---

[19] We chose total assets as a deflator rather than market value so that our measure is unaffected by stock price movements. Results are consistent deflating by market value.

[20] This is not to suggest that U.S. firms necessarily manage earnings less than those in other countries since, as argued in Brown and Higgins (2001), they may be managing earnings to meet analyst forecasts.

TABLE 7
Accrual/Cash Flow Correlations

Country	*Subperiod*	*Coefficient*	*Z-Statistic*
Australia	First	−0.790	
	Second	−0.707	−4.92
Canada	First	−0.655	
	Second	−0.594	−0.32
Germany	First	−0.919	
	Second	−0.868	−3.52
France	First	−0.897	
	Second	−0.886	−2.36
U.K.	First	−0.795	
	Second	−0.686	−7.92
Japan	First	−0.940	
	Second	−0.835	−11.73
U.S.	First	−0.557	
	Second	−0.565	1.86

Results are rank correlations of accruals and cash flows from operations, both deflated by total assets. The first subperiod is 1987–1992 and the second is 1994–1999. The Z-statistic is from a rank sum test of the significance of the change in correlations between periods.

al. (2002), we test for the significance of differences by estimating by-firm correlations in each subperiod and conducting a rank sum test across the two periods. There is a significant decrease (at the .01 level) in the correlation magnitude for each country except the U.S. and Canada.[21] The fact that evidence of a change is weakest in the U.S. and Canada, where the pricing evidence was also weakest, suggests that the two sets of tests may reflect the same underlying phenomenon. The U.S. is the only country that experienced a more negative correlation in the second period, although the difference is small (−.565 vs. −.557, Z-statistic = 1.86). Our results also suggest that predictable differences remain across countries, with Germany, Japan, and France showing the greatest evidence of smoothing, even in the second subperiod.[22]

V. OTHER ANALYSIS

Our primary focus is on earnings since this measure has attracted the most attention from regulators and analysts, and it should also be more sensitive than other performance measures to accounting changes in the short term. The balance sheet, on the other hand, would not be expected to change quickly in response to changes in accounting because shareholders' equity includes the effects of accounting choices made over the life of the firm and accounting changes are generally applied prospectively. Nevertheless, if changes in accounting have made measurement more comparable, then we would expect to see a narrowing of differences in balance sheet valuation over time.

Table 1 presents statistics on book-to-market ratios over the sample period. Following our earnings analysis, we regress mean-adjusted book-to-market ratios on the 14 country/subperiod indicators. Again, we base our inference on whether coefficient estimates move closer to the mean in absolute value. Results (not tabulated) provide evidence of convergence, although it is not as strong as the evidence for earnings/price ratios. In the second period, cross-country differences were lower for five of the seven countries. Computing a rank sum statistic for the first half vs. the second half, the reduction in variance of medians is significant at the .05 level. Similarly, examining the relation between dispersion of book/market ratios and time, there is a significant negative correlation over our sample period. Including GDP growth, sales growth, EPS growth, and real interest rates has no effect on the inference. Overall, there is some evidence of convergence in book-to-market ratios, but it is not as consistent as for earnings.

VI. CONCLUSIONS

Our results suggest that earnings valuation across countries has become more similar over time. Results are robust to controls for underlying economic activity and are generally consistent across sample countries, but are strongest for Germany and Japan. The changes in multiples are robust to controls for the pricing of cash flows, suggesting that the effects are driven by the pricing of accruals, consistent with an increase in comparability of accounting over time. That conclusion is reinforced by evidence that the multiples applied to accruals have converged. However, significant predictable differences remain suggesting that, while comparability appears to have improved, the process is incomplete.

---

[21] These results are generally consistent with those in Ball, Kothari, and Robin (2000), which focuses on differences in conservatism (which they interpret as evidence of smoothing) across countries. In supplemental analysis, they find evidence of reduced conservatism for Australia, Germany, France, and U.K., but not for the U.S., Canada, or Japan. Bhattacharya et al. (2001) also compare earnings management over time using different measures and provide evidence that it is generally declining.

[22] We also examined changes in the explanatory power of earnings for returns over time, but found no clear pattern across the two subperiods. However, we were unable to document evidence of systematic cross-country differences in either subperiod, suggesting that this approach may lack power.

Results on the correlation between accruals and cash flows also indicate convergence, consistent with the conclusions based on the pricing multiples. Japan, Germany, and France show the most negative relation between cash flows and accruals, while the U.S. and Canada are the least negative. Again, evidence of a change is clearest for Japan and Germany.

While one cannot draw strong policy implications based on evidence of this type, our results raise the possibility that, even in the face of fairly haphazard global accounting standard development, accounting practices may be becoming more similar. Given the implicit and explicit pressures toward harmonization of financial reporting, it is perhaps not surprising that differences are declining. However, the results also suggest that the effects of traditional differences, while reduced, still persist.

The analysis is limited in many ways and is subject to numerous caveats. First, our conclusions are only as good as our methods. We apply approaches that have been used in the literature, attempt to control for as many variables as possible and consider a variety of analyses with similar conclusions. Nevertheless, we are unable to control for all potential omitted variables, so our results should be viewed with caution. Further, our sample is limited to the largest firms in the largest economies. Many of the more difficult regulatory issues relate to small young firms in emerging industries and less-developed markets where financial reporting is less developed and enforcement is limited.

REFERENCES

Accountancy. 1997a. French reforms going through. (February): 12.

———. 1997b. U.K. businesses feel IASC pressure. (March): 8.

Alexander, D., and S. Archer. 1991. *European Accounting Guide*. San Diego, CA: HBJ Professional Publishing.

Alford, A., J. Jones, R. Leftwich, and M. Zmijewski. 1993. The relative information content of accounting disclosures in different countries. *Journal of Accounting Research* 31 (Supplement): 183–223.

Ball, R., S. P. Kothari, and A. Robin. 2000. The effect of international institutional factors on properties of accounting earnings. *Journal of Accounting and Economics* 29 (October): 1–51.

Ball, R., A. Robin, and J. S. Wu. 2000. Incentives versus standards: Properties of accounting income in four East Asian countries, and implications for acceptance of IAS. Working paper, University of Rochester.

Barth, M., and G. Clinch. 1996. International accounting differences and their relation to share prices: Evidence from U.K., Australian, and Canadian firms. *Contemporary Accounting Research* 13 (Spring): 135–170.

Beatty, A., B. Ke, and K. Petroni. 2002. Earnings management to avoid earnings declines across publicly and privately held banks. *The Accounting Review* 77 (3): 547–570.

Bhattacharya, U., H. Daouk, and M. Welker. 2001. The world price of earnings opacity. Working paper, Indiana University.

Biener, H. 1994. What is the future of mutual recognition of financial statements and is comparability really necessary? *European Accounting Review* 3: 335–342.

Brown, L., and H. Higgins. 2001. Managing earnings surprises in the U.S. versus 12 other countries. *Journal of Accounting and Public Policy* 20 (Winter): 373–398.

Burgstahler, D., and I. Dichev. 1997. Earnings management to avoid earnings decreases and losses. *Journal of Accounting and Economics* 24 (December): 99–126.

CA Magazine. 1996. Champion of Canadian standards. (January/February): 10.

Cairns, D. 1997. The future shape of harmonization: A reply. *European Accounting Review* 6 (2): 305–348.

Choi, F. D. S., C. Frost, and G. Meek. 1999. *International Accounting*. 3rd edition. Upper Saddle River, NJ: Prentice Hall.

Cochrane, J. 1994. Are U.S. regulatory requirements for foreign firms appropriate? *Fordham International Law Journal* 17: 58–67.

———, J. Shapiro, and J. Tobin. 1996. Foreign equities and U.S. investors: Breaking down the barriers separating supply and demand. *Stanford Journal of Law, Business and Finance* (Summer): 241–263.

Curran, B. 1996. An Australian comparison. *Accountancy* (September): 64–65.

Dhaliwal, D., D. Guenther, and M. Trombley. 1999. Inventory accounting methods and earnings-price ratios. *Contemporary Accounting Research* (Fall): 419–436.

Doupnik, T. 1992. Recent innovations in German accounting practice through the integration of EC directives. *Advances in International Accounting* 5: 75–103.

Economist. 1999. Rupert laid bare. (March 20): 63–64.

Financial Accounting Standards Board (FASB). 1996. *The IASC-U.S. Comparison Project: A Report on the Similarities and Differences between IASC Standards and U.S. GAAP*. Edited by C. Bloomer. Stamford, CT: FASB.

French, K. R., and J. M. Poterba. 1991. Were Japanese stock prices too high? *Journal of Financial Economics* 29 (October): 337–362.

Frost, C., and K. Ramin. 1997. Corporate financial disclosure: A global assessment. In *International Accounting and Finance Handbook*, edited by F. D. S. Choi. New York, NY: John Wiley & Sons.

Guenther, D., and D. Young. 2000. The association between financial accounting measures and real economic activity: A multinational study. *Journal of Accounting and Economics* 29 (February): 53–72.

Jenkins, E. 1999. Financial reporting in a global capital market world. *Financial Accounting Series* 198-A (June 29): 2–6.

Joos, P., and M. Lang. 1994. The effects of accounting diversity: Evidence from the European Union. *Journal of Accounting Research* 32 (Supplement): 141–175.

Lang, M., J. Raedy, and M. Yetman. 2002. How representative are cross-listed firms? An analysis of firm performance and accounting quality. Working paper, The University of North Carolina at Chapel Hill.

Leuz, C., D. J. Nanda, and P. Wysocki. 2001. Investor protection and earnings management: An international comparison. Working paper, University of Pennsylvania.

Myers, L., and D. Skinner. 2001. Earnings momentum and earnings management. Working paper, University of Michigan.

Pape, J. 1997. A milestone in German accounting. *Accountancy* 119 (May): 58–59.

Pownall, G., and K. Schipper. 1999. Implications of accounting research for the SEC's consideration of International Accounting Standards for U.S. securities offerings. *Accounting Horizons* 13 (September): 259–280.

Shibata, Y. 1998. Turnabout in Japan. *Global Finance* 12 (November): 44–45.

Simmonds, A., and O. Azieres. 1989. *Harmonisatie van de verslaggeving in Europa: Voltooid in het jaar 2000*. London, U.K.: Touche Ross Europe.

Speidell, L. S., and V. B. Bavishi. 1992. GAAP arbitrage: Valuation opportunities in International Accounting Standards. *Financial Analysts Journal* (November/December): 58–66.

THE ACCOUNTING REVIEW
Vol. 77
Supplement 2002
pp. 135–138

DISCUSSION OF Empirical Evidence on the Evolution of International Earnings

David A. Guenther
University of Colorado at Boulder

I. INTRODUCTION

In "Empirical Evidence on the Evolution of International Earnings" Land and Lang (2002) (hereafter the paper) investigate a question related to international accounting differences: Has there been convergence of earnings multiples over time, supporting an inference that financial accounting measures used in different countries become more similar over time? The authors make a convincing argument that anecdotal evidence would lead us to expect such a convergence in accounting measures; and although their research design does not allow them to make a causal link between their empirical evidence and accounting changes across countries, they do a thorough job of attempting to control for alternate explanations. After reading the paper the reader is left with two strong impressions. First, average earnings/price (E/P) ratios across the countries analyzed in this paper (Australia, Canada, France, Germany, Japan, the U.K., and the U.S.) have moved closer together from the 1987–1992 time period to the 1994–1999 time period. Second, this convergence of E/P ratios seems to be related to accounting accruals.

Whether E/P ratios have converged over time is an empirical question, as the authors point out, and it seems appropriate that the question be studied by examining the data as the authors have done. The authors use different types of tests to try to rule out alternate explanations for the convergence of E/P ratios, and the results of these tests seem consistent with expectations given the anecdotal evidence cited by the authors (i.e., the results are not surprising).

As noted by some of the Conference participants, the apparent convergence of financial reporting across countries is evidence only of the convergence of accounting *measurement* since it is based on reported net income. Convergence of accounting measurement would suggest that net income is being computed using similar accounting methods across countries. An example of convergence in measurement is the use of the comprehensive (rather than the partial) method of recording deferred tax liabilities, since this affects reported deferred tax expense and, hence, reported net income. Another question of equal interest to researchers is whether there has been a convergence in accounting *disclosure* over this same time period. An example of convergence in disclosure is disclosing the market value

---

Helpful comments were received from Katherine Schipper and Mark H. Lang.

of all financial assets and liabilities, since this does not affect reported net income. If capital markets are efficient with respect to accounting information, then disclosure of value-relevant information will affect stock prices, and therefore the convergence of accounting disclosure should also be of interest to accounting researchers.

Unfortunately the research design of this study cannot address the question of whether accounting disclosures have also converged. Also, the results of this study cannot tell us anything about what has happened to accounting standards in different countries over this time period, since the net income measure used in the study may reflect choices made by firms to report using either U.S. GAAP or International Accounting Standards (IAS) rather than the GAAP of their home country; that is, the convergence of E/P ratios reflects a convergence toward using either U.S. GAAP or IAS and not a convergence of standards, per se. The authors report that the proportion of sample firms reporting using IAS or U.S. GAAP was 0.8 percent in the first period and 2.1 percent in the second period, and that the conclusions are not sensitive to excluding these observations. Therefore, although a shift to IAS or U.S. GAAP appears to have occurred during the sample period, such a shift does not affect the paper's results.

II. IMPACT OF STOCK MARKET INDEX INCREASES

One alternative explanation for the convergence in E/P ratios documented by the paper is that the convergence is due to denominator effects, not numerator effects. That is, the results might reflect the increases in stock market indexes that occurred in all of the sample countries (except Japan) over the period 1987–1992 to 1994–1999. Since the dependent variable in these tests is the ratio of earnings to stock price, systematic increases in stock prices in the sample countries would decrease average E/P ratios. Similarly, systematic decreases in stock prices in Japan over this same time period would increase average E/P ratios for the Japanese firms in the sample.

Although the authors argue that their results suggest E/P ratios converge over their sample time period, their results can also be interpreted as demonstrating that E/P ratios *decreased* over the time period in five of the seven sample countries (Australia, U.K., France, U.S., and Canada), while E/P ratios *increased* in two of the sample countries (Germany and Japan). Changes in stock market indexes over the same time period can explain these changes in six out of the seven sample countries, with only the results for Germany being inconsistent with a stock market-based explanation. To test this explanation, for each of the sample countries I compute the percentage change in the country's major stock market index from the January 1, 1990 level to the January 1, 1997 level. I also used data from Land and Lang (2002, Table 1) to compute the percentage change in mean E/P ratios across the two time periods. These results are shown in Table 1.

When Germany is excluded from the sample countries, the Pearson correlation between the percentage change in E/P ratios and the percentage change in market indexes is –0.83. A regression of the change in E/P ratio on the change in market index has an R^2 of 0.70. Thus it appears that, except for Germany, a large part of the change in E/P ratios observed over this time period may be related to changes in stock prices in the sample countries.

Table 1 also shows the mean cash flow/price ratios for the sample countries for the two subperiods (taken from Land and Lang [2002, Table 3]), as well as the percentage changes in these ratios. It is apparent from these data that the pattern observed for the E/P ratios (i.e., decreases in ratios for five of the sample countries and increases in ratios for Germany and Japan) is also present in the cash flow/price ratios. The Pearson correlation between percentage changes in E/P ratios and percentage changes in cash flow/price ratios is 0.97. This suggests that the pattern observed in the E/P ratio changes is also

TABLE 1

Changes in E/P Ratios, Stock Market Indexes, and Cash Flow/Price Ratios from 1987–1992 to 1994–1999

	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*	*(8)*	*(9)*
	E/P Ratio 1987–1992	*E/P Ratio 1994–1999*	*(2)/(1)*	*Market Index 1990*	*Market Index 1997*	*(5)/(4)*	*Cash Flow/ Price Ratio 1987–1992*	*Cash Flow/ Price Ratio 1994–1999*	*(8)/(7)*
Australia	0.095	0.072	0.753	1,655	2,411	1.457	0.143	0.121	0.842
Canada	0.070	0.066	0.938	4,004	5,904	1.475	0.118	0.115	0.971
Germany	0.042	0.068	1.611	1,443	2,849	1.974	0.090	0.117	1.297
France	0.080	0.066	0.826	1,832	2,257	1.232	0.128	0.115	0.899
U.K.	0.089	0.077	0.871	2,434	4,057	1.667	0.137	0.126	0.923
Japan	0.019	0.038	1.938	38,713	19,446	0.502	0.067	0.087	1.285
U.S.	0.073	0.061	0.841	360	737	2.047	0.121	0.110	0.912

Stock market indexes are All Ordinaries (Australia), TSE 300 Composite (Canada), DAX (Germany), CAC 40 (France), FTSE 100 (U.K.), Nikkei 225 (Japan), and S&P 500 (U.S.).

E/P ratios are from Land and Lang (2002, Table 1). Market indexes are from Yahoo! Finance. Market indexes for Germany and France were not available as of January 1, 1990, so the earliest date in 1990 for which data were available was used for those countries. Cash flow/price ratios are from Land and Lang (2002, Table 3).

present in the cash flow/price ratio changes, and that, except for Germany, this pattern can be explained by changes in stock market indexes in the sample countries.

Obviously, the results for Germany are not consistent with the stock index explanation, since the German stock market index increased over the sample period, while both E/P ratios and cash flow/price ratios for German firms also increased. However, regardless of the cause of the increase in E/P ratios for German firms over the sample period, the same increase appears in the cash flow/price ratios as well. The difference is that the increase in E/P ratios causes German firms to move closer to the mean, while the increase in cash flow/price ratios causes German firms to move further from the mean.

III. ALTERNATE EXPLANATIONS

Although the authors have been careful to design tests to rule out several competing explanations for their results, some alternate explanations suggested by conference participants warrant mention in this discussion. First, the way investors use accounting information to set stock prices may be changing over time. For example, increasing sophistication on the part of investors in the sample firms would be expected to change the observed relation between reported accounting earnings and stock prices. This might occur if the sample firms are large multinationals of the type that mutual fund managers are likely to follow, and follow increasingly over the sample period considered. Second, the integration of global capital markets over this time period could lead to consistent pricing of risks in all markets. Third, several Conference participants thought that the results were simply an example of mean reversion over time.

The authors have done a good job of addressing alternate explanations by including several control variables in their tests. However, the use of these control variables also led to comments by Conference participants. The way control variables are used in the paper assumes that the effect of a particular control variable remains constant over time, but the level of the control variable changes. Conference participants suggested that an alternative approach would be to allow the effect of the control variable (as measured by the regression coefficient) to change over time as well. This would require the inclusion of interactions between the time period indicator variables and the control variables. The authors have included such an interaction between the cash flow/price ratios and the country and subperiod indicator variables in the paper's Table 5 and Table 6 regressions.

IV. CONCLUSION

The results reported in the study are difficult to interpret for two reasons. First, both changes in E/P ratios and changes in cash flow/price ratios exhibit a similar pattern over time, in that both sets of ratios are decreasing for five of the sample countries (Australia, Canada, France, U.K., and U.S.), and both sets of ratios are increasing for Germany and Japan. For all of the countries except Germany, this result can be explained by changes in stock market indexes over the same time period. I interpret this as evidence that whatever is causing increases or decreases in E/P ratios is also causing similar increases and decreases in cash flow/price ratios. In my view, the reader is left to wonder whether the apparent convergence reported in the paper can be explained by stock price movements for six of the sample countries, and an anomalous result for the German firms.

The authors take a different view of the cash flow/price results, maintaining that, although there are significant changes in some cash flow/price ratios over the period, there is not a consistent pattern, since four countries' multiples move away from the average. In reaching their conclusion that there is little similarity between the movements of E/P and cash flow/price ratios the authors are comparing movement toward or away from the mean. However, if the comparison is between increases or decreases in ratios, then the pattern is nearly identical between the two sets of measures.

The second reason the results are difficult to interpret is that, even if the authors are correct and there has actually been convergence in E/P ratios across the sample countries over time, what should we make of this result? Absent specific knowledge of the cause of the convergence, it is difficult to draw public policy or regulatory implications. We do not know if organizations such as the IASC have been successful in their drive to harmonize accounting standards, whether behavior of managers has changed, or whether demand by capital markets has led to the use of similar accounting measures. Without more understanding of what is behind the convergence of E/P ratios, the results in and of themselves leave the reader wondering what conclusions to draw from the paper.

REFERENCE

Land, J., and M. H. Lang. 2002. Empirical evidence on the evolution of international earnings. *The Accounting Review* (Supplement): 115–133.

THE ACCOUNTING REVIEW
Vol. 77
Supplement 2002
pp. 139–167

Evaluating Financial Reporting Quality: The Effects of Financial Expertise vs. Financial Literacy

Linda McDaniel
University of Kentucky
Roger D. Martin
Laureen A. Maines
Indiana University

ABSTRACT: Audit committees evaluate financial reporting quality as part of their corporate oversight responsibilities. Given this responsibility, the national stock exchanges now require all audit committee members to be financially literate and at least one member to have financial expertise. In light of recent debates over this requirement, we provide evidence on how experts and literates differ in their evaluations of financial reporting quality. Results suggest that experts' evaluations of financial reporting quality are more strongly associated with their assessments of characteristics underlying reporting quality (e.g., relevance) espoused in Statement of Financial Accounting Concepts No. 2's framework than literates' evaluations. Additionally, literates are more likely than experts to identify concerns about reporting treatments for business activities that are prominent in the business press or are distinguished by their nonrecurring nature, while experts are more likely to raise concerns about reporting treatments for less prominent, recurring activities. This same pattern occurs in the ratings of the quality of the reporting treatments for specific financial statement items with respect to elements underlying reporting quality (e.g., neutrality); literates (experts) assess the quality elements for the reporting treatments of prominent and nonrecurring items (less prominent and recurring items) comparatively lower than experts (literates). These results suggest that including financial experts on audit committees is likely to change the structure

---

We thank Walt Blacconiere, Frank Hodge, Pat Hopkins, Michael Kimbrough, Bill Kinney, Bernadine Low, Elaine Maudlin, Mark Nelson, Mark Peecher, Fred Phillips, Bob Ramsey, Jerry Salamon, Geoff Sprinkle, Dan Stone, Nathan Stuart, and Greg Trompeter for written comments, and participants at Boston College, Emory University, Indiana University, University of Kentucky, and University of Missouri research workshops and *The Accounting Review*'s Conference on the Quality of Earnings for comments on earlier versions of the paper. Additionally, we have benefited from discussions with Jeff Brovet, Joe Carcello, Robert Dunaway, Ed Evans, Dave Greene, John Hassell, Bill Kinney, Ganesh Krishnamoorthy, David Landsittel, Bob Libby, and Rob Parry, and the research assistance of Andrea Astill, Michael Kimbrough, Kelly McGaffrey, Craig Nichols, Dan Phan, and David Upton. Finally, we thank Ann Gibson, Larry Scott, and Art Wachholz for help in contacting experimental participants and the participants for their time and effort in completing the research study.

and focus of audit committee discussions about financial reporting quality, and may affect the committee's overall assessment of the quality of a company's financial reports.

Keywords: *audit committees; financial expertise; financial literacy; financial reporting quality.*

Data Availability: *Contact the authors.*

I. INTRODUCTION

We investigate the extent to which financial experts make judgments about financial reporting quality that differ from those of financial literates. Given that the national stock exchanges and the National Association of Securities Dealers (NASD) now require financial literacy for all audit committee members and financial expertise for at least one member (PricewaterhouseCoopers 2000a, 5), examining differences between experts' and literates' judgments is a first step in understanding how the expertise requirement may affect audit committees' behaviors.[1]

While most see financial literacy as a necessary requirement to be an effective audit committee member (Association for Investment Management and Research [AIMR] 1999; Ernst & Young 2000), there is disagreement as to whether the financial expertise requirement is justified. For example, some have expressed concerns about the difficulty of finding qualified audit committee members (Kirk 2000; Olson 1999), because of increased legal exposure for designated experts (Paskell-Mede and Jackson 1999; Zacharias 2000). Perhaps a more critical argument is that audit committee members with financial expertise may actually be less effective than members with other characteristics, such as practical management experience (AIMR 1999; Olson 1999). Thus, evidence is needed on the likely effects of the financial expertise requirement on audit committee behavior. One possible effect could stem from differences in experts' vs. literates' evaluations of financial reporting quality.

We investigate whether and how financial experts' judgments related to financial reporting quality differ from those of financial literates in an experiment in which financial experts (audit managers) and financial literates (recent Executive M.B.A. graduates) assume the roles of audit committee members. Participants are asked to evaluate the financial reporting quality of a 10-Q report for a fictitious textile company in preparation for the mandatory quarterly discussion about financial reporting quality with auditors (AICPA 1999). Audit managers and Executive M.B.A. students meet the stock exchanges' definitions of financial experts and financial literates, respectively. We use these two groups as participants rather than individuals currently serving on an audit committee to isolate the effects of differences in financial knowledge on financial reporting quality judgments.[2] Additionally, in our experiment, we use features of the Financial Accounting Standard Board's conceptual framework, as reported in Statement of Accounting Concepts No. 2 (SFAC No. 2), *Qualitative Characteristics of Accounting Information* (Financial Accounting Standards Board [FASB] 1980), as an appropriate framework for evaluating reporting quality.

---

[1] Financial literacy is generally described as the ability to read and understand basic financial statements (Blue Ribbon Committee 1999, 26), while financial expertise typically is framed in terms of employment experience or certification in accounting/finance (PricewaterhouseCoopers 2000a, 4).

[2] We sought participants who differed in financial knowledge, but had similar levels of overall business-related and industry-specific knowledge. In particular, we wanted participants to have a significant amount of business experience (to qualify as potential audit committee members), but low industry experience (enabling us to separate effects of financial knowledge from effects of industry knowledge). Based on discussions with various individuals who work with audit committees, we believed that using actual audit committee members would reduce our ability to isolate the effects of financial knowledge, as opposed to other types of knowledge, including industry knowledge. However, by choosing only individuals with low industry experience, we are assuming that the influence of industry experience on individuals' evaluations of financial reporting quality does not interact with their financial knowledge.

We examine two aspects of an audit committee's evaluation of financial reporting quality. First, given an audit committee's responsibility to discuss the overall quality of a company's financial reporting (SEC 1999, 2), we test for differences in experts' and literates' frameworks for evaluating reporting quality by examining the relations between assessments of overall reporting quality and assessments of three quality characteristics (relevance, reliability, and comparability) taken from SFAC No. 2. We predict that financial experts' evaluations of overall financial reporting quality will more strongly reflect their assessments of these three quality characteristics than will literates' evaluations. Specifically, we examine whether experts' assessments for the three individual quality characteristics better aggregate to the overall financial reporting quality assessment than do literates' assessments.

Second, given that audit committees address specific reporting-quality concerns about individual financial statement items in their discussions with auditors (AICPA 1999, para. 11), we examine differences in experts' and literates' identifications of reporting concerns/issues and in their evaluations of the quality of the reporting treatment for individual financial statement items. Based on their reviews of the company's financial statements, participants identified and ranked three reporting issues they deemed most critical for discussion with the auditors. Subsequently, they evaluated the quality of the reporting treatment for ten specified financial statement items with respect to selected elements of reporting quality espoused in SFAC No. 2, i.e., feedback and predictive value for the relevance characteristic, and verifiability, comprehensiveness, representational faithfulness, and neutrality for the reliability characteristic.

The reporting treatment for each of the ten specified financial statement items had potential implications for the quality of the financial statements.[3] We classified each financial item into one of three salience categories (representing a within-subjects factor) based on a combination of two attributes. The two attributes related to whether the item (1) had recently received low vs. high prominence in the business press and (2) reflected a recurring, routine business activity vs. a nonrecurring, unusual business activity. We expect that literates will identify reporting concerns primarily based on an item's salience, created by its public prominence and/or nonrecurring or unusual nature (e.g., business reorganization). In contrast, we expect experts to be less influenced by an item's salience due to their broader knowledge base and greater exposure to recurring and nonrecurring business activities. Thus, we predict literates are more likely than experts to identify concerns related to the financial reporting treatment of publicly prominent, nonrecurring financial statement items, while experts are more likely than literates to raise issues about the financial reporting treatment of low prominence, recurring items. We expect that literates' and experts' ratings of the quality of the reporting treatment for specific financial items with respect to elements underlying reporting-quality characteristics (e.g., neutrality) will follow similar patterns.

Our results suggest that financial experts' frameworks for evaluating overall financial reporting quality for a set of financial statements differ from those of financial literates. Specifically, experts' individual assessments of the relevance and comparability characteristics of quality espoused in SFAC No. 2 better aggregate to their overall assessments of reporting quality, while literates' evaluations of overall reporting quality were unrelated to their assessments of relevance and comparability. Neither group's overall reporting-quality assessments reflected the reliability assessment.

---

[3] We interpret the term "financial statement items" broadly to include not only items reported within the body of the financial statements, but also items reported in the footnotes and MDA. Also, we use the terms "financial items" or "items" synonymously with "financial statement items."

Our results also indicate that literates were more likely than experts to raise primary concerns about the reporting treatments for high-salience financial statement items, i.e., items prominent in the business press or items distinguished by their unusual, nonrecurring nature. Literates also rated underlying elements of the quality of reporting treatments of high-salience items lower than did experts. In contrast, experts were more likely than literates to identify reporting concerns related to recurring business activities that have received little business press coverage and provided lower assessments of the underlying quality of the reporting treatment of such items than did literates.

This study identifies several differences between the judgments of financial experts and financial literates. Assuming that the individual differences found in our study extrapolate to judgments made by real audit committees in a group setting, our results form a preliminary basis for assessing how the financial expertise requirement will affect audit committee behavior. Our findings indicate that financial experts possess a framework for evaluating financial reporting quality that is more consistent with SFAC No. 2 and, thus, are likely to bring more structure to discussions of reporting quality compared to literates. These results suggest that experts and literates are unlikely to have a common basis for discussing reporting issues, and highlight the need either to educate literates on the SFAC No. 2 framework or to develop a new, shared framework for evaluating financial reporting quality, as recommended by Jonas and Blanchet (2000, 358).

Our findings also indicate that experts and literates bring differing viewpoints to the identification and evaluation of specific financial reporting concerns. Experts view as most important financial reporting concerns related to less prominent and recurring business activities while literates initially focus on concerns about nonrecurring, prominent (in the press) financial items. In addition, our results suggest that financial items that are less critical to overall reporting quality can distract literates. Overall, our findings suggest that the presence of experts may both sharpen and shift the focus of audit committees' discussions and overall evaluations of a company's financial reporting quality. These changes likely will improve audit committees' evaluations of financial reporting quality since more time and attention will be devoted to recurring issues important to financial reporting quality.

The remainder of the paper is organized as follows. Section II summarizes the background of recent activities and rule making related to audit committees and financial reporting quality, discusses relevant research, and develops the hypotheses. Section III describes the experiment, Section IV presents results, and Section V discusses implications of these results and concludes the paper.

II. BACKGROUND AND HYPOTHESES DEVELOPMENT

Improving Audit Committee Oversight Effectiveness

Audit committee members (as representatives of the board of directors), along with management and auditors provide oversight with respect to financial reporting (Levitt 2000). Audit committees have specific responsibility for overseeing financial reporting activities (Braiotta 1999, 104). Regulators and the accounting profession have touted the role of audit committees in protecting investors, and accounting research suggests that market participants see audit committees as providing meaningful oversight of the financial reporting process (DeFond and Jiambalvo 1991; Dechow et al. 1996; McMullen 1996; Wild 1996).[4]

Responding to the SEC's concerns about erosion in the quality of financial reporting (Levitt 1998), the NYSE and NASD sponsored the Blue Ribbon Committee on Improving the Effectiveness of Corporate Audit Committees (Blue Ribbon Committee), which issued ten recommendations aimed at improving the oversight effectiveness of audit committees and promoting quality financial reporting (Blue Ribbon Committee 1999, 3).

The Auditing Standards Board and the SEC recently adopted one recommendation of the Blue Ribbon Committee that requires auditors to discuss on a quarterly basis the quality of financial reporting with the audit committee (AICPA 1999; SEC 1999, 2). Additionally, the New York Stock Exchange (NYSE), American Stock Exchange (AMEX), and National Association of Securities Dealers (NASD) adopted a recommendation that all audit committee members be financially literate and at least one member demonstrate a higher level of financial reporting knowledge, referred to as financial expertise (Blue Ribbon Committee 1999, 25; PricewaterhouseCoopers 2000a, 5). Financial literacy is described as the ability to read and understand fundamental financial statements (Blue Ribbon Committee 1999, 26), although the NYSE's rules allow each board of directors to interpret this term.[5] Financial expertise is defined somewhat differently by the exchanges; however, all define financial expertise as related to past employment experience in finance/accounting or professional certification in accounting or finance (PricewaterhouseCoopers 2000a, 4).[6]

While few have questioned the merits of requiring auditors and audit committee members to discuss the quality of financial reporting, the benefits of the expertise qualification have been debated. The expertise requirement assumes that including at least one financially literate audit committee member will improve the committee's effectiveness, for example, by leading audit committee members to identify and ask questions that "make management think harder and auditors dig deeper" (Levitt 2000). Yet, even the Blue Ribbon Committee acknowledged, "a director's ability to ask and intelligently evaluate answers to such questions may not require expertise" (Blue Ribbon Committee 1999, 25). Similarly, the AIMR has taken the position that a background in accounting or finance should not be "necessary and sufficient" for qualification, and points out that individuals without formal training "can be as insightful (or in some cases more so)" than individuals with formal financial training (AIMR 1999).

---

[4] For example, for over 60 years the SEC has recommended that companies form audit committees (e.g., Accounting Series Release (ASR) No. 19, SEC 1940). The NYSE instituted audit committee requirements for its listed companies in 1978. The accounting profession and several special commissions (e.g., the Public Oversight Board (POB) Advisory Panel on Auditor Independence (1994), i.e., the "Kirk Panel") have addressed the role of audit committees in corporate governance and the enhancement of the quality of financial reporting.

[5] The literacy requirements appear to be minimal. For example, an inspection of Costales and Szurovy (1994), one of the financial literacy resources suggested by Ernst & Young (2000), indicates a focus on understanding the three basic financial statements, with little discussion of specific applications of GAAP.

[6] Specifically, the NYSE requires one audit committee member to have accounting or related financial management expertise, while the NASD/AMEX frames its requirement in terms of past employment experience in finance or accounting, certification in accounting, or other comparable experience resulting in the individual's financial sophistication (PricewaterhouseCoopers 2000a, 4).

To initiate an inquiry into whether financial expertise improves an audit committee's effectiveness at evaluating financial reporting quality, we examine whether and how financial experts and literates differ in their evaluations of selected aspects of reporting quality.[7] Because the Blue Ribbon Committee and national stock exchanges have provided little guidance for predicting whether and how experts and literates will differ in this regard, we rely on previous academic research to guide our predictions.

Financial Expertise in Evaluating Financial Reporting Quality

The judgment and decision-making literature provides evidence that experts and nonexperts differ in their representations of judgment and decision situations (Yates 1990, 372). These differences are attributed primarily to knowledge differences that relate to either specific instances stored in memory (episodic knowledge) or abstracted general properties (semantic knowledge) (Anderson 2000, 164). Abstracted semantic knowledge holds across many instances and is often referred to as a mental model or schema (Fiske and Taylor 1991, 98). Individuals develop schemas by using experiences to form associations among instances/episodes into structured, meaningful patterns or categories (Castellan 1993, 226). In evaluating decision situations, schemas are activated (and strengthened) by explicit thought about a topic or an encounter with information relevant to the schema. Once activated, schemas influence memory retrieval, direct attention to specific types of information, and affect the decision representation and interpretation of this information (Smith 1998, 403).

We propose that financial experts' and literates' semantic and episodic knowledge about financial reporting quality differ as follows. First, we assume experts' episodic knowledge about reporting quality represents both first-hand experiences with relevant problems and second-hand experiences (e.g., discussions with other experts). In contrast, literates' episodic knowledge likely is based more on second-hand information, such as accounting events reported in the business press. Second, we assume experts' and literates' semantic knowledge (and thus their problem representations) will differ due to the number and nature of first- and second-hand encounters with reporting-quality concerns (see e.g., Libby 1995). We predict that these knowledge differences will affect the way in which experts vs. literates (1) assess overall financial reporting quality and incorporate underlying characteristics of reporting quality into these assessments and (2) identify and evaluate critical reporting issues. Specifically, we predict that experts will possess better-developed schemas or frameworks for evaluating reporting quality relative to literates. In addition, due to the absence of an experienced-based schema and limited second-hand episodic knowledge, we expect that, relative to experts, literates' identifications of critical reporting issues and evaluations of the quality of reporting treatments will be more influenced by attributes (i.e., salience) of the presented information (see, e.g., Tversky and Kahneman 1973). We discuss each of these in turn.

Framework for Evaluating Overall Financial Reporting Quality

The SEC, auditing profession, and national exchanges have not specified an explicit definition of or a framework for financial reporting quality (see e.g., Blue Ribbon

---

[7] Two recent studies have addressed other issues related to expertise in audit committee settings. DeZoort and Salterio (2001) find that audit committee members who have more financial knowledge are more likely to support auditors than management in a "substance over form" revenue recognition policy dispute. DeZoort (1998) shows that more experienced audit committee members assess internal controls as stronger than less experienced members. Neither study directly examines audit committee members' evaluations of financial reporting quality.

Committee 1999, 2–3; AICPA 2000; Turner 2001).[8] One of these groups (the Auditing Standards Board) asked the FASB for guidance on defining financial reporting quality and the FASB referred to SFAC No. 2 as an accepted framework (Jenkins 2000). We assume the framework in SFAC No. 2 is an appropriate model for an audit committee to apply in its task of evaluating the quality of a firm's reporting, and we use this framework to examine how experts and literates differ in their assessments of overall financial reporting quality (FASB 1980). Specifically, we use SFAC No. 2's characteristics of relevance, reliability, and comparability to capture characteristics related to overall financial reporting quality.

We assume that financial experts, due to their required financial-related experiences or certifications, have developed schemas for reporting quality consistent with the SFAC No. 2 framework. For example, the certification requirement (CPA or equivalent) implies study that would lead to direct knowledge of SFAC No. 2. Also, experts' knowledge (e.g., of the qualitative characteristics and their association with overall financial reporting quality) develops through experience (application of authoritative literature) in evaluating financial statements with these characteristics in mind (see e.g., Libby and Luft 1993).

We assume that individuals designated as financially literate will have some familiarity with reporting quality as described in SFAC No. 2 from second-hand exposures to the popular press, prior or continuing financial education, or discussions in meetings.[9] However, this familiarity is likely to be more limited and to erode over time since literates do not reinforce and/or further develop the framework via practical experience. Overall, because they lack the financial certification or experience necessary to be designated a financial expert, literates will have a limited ability to form and access a well-defined financial reporting quality schema that delineates how specific characteristics are associated with overall financial reporting quality for a set of financial statements.

These arguments suggest that experts' frameworks for evaluating financial reporting quality will be better developed with respect to SFAC No. 2 quality characteristics (relevance, reliability, and comparability) than those of financial literates. This leads to the following alternative-form hypothesis:

> **H1:** Financial experts' evaluations of financial reporting quality are more strongly associated with their assessments of characteristics underlying reporting quality espoused in SFAC No. 2's framework than literates' evaluations.

Identification of Critical Reporting Issues and Evaluations of the Elements of Reporting Quality for the Reporting Treatment of Specific Financial Items

While auditors and audit committees are charged with evaluating overall reporting quality, actual discussions about quality typically focus on individual financial statement items (e.g., PricewaterhouseCoopers 2000b; AICPA 2000). Such discussions likely focus on possible concerns about specific elements of the characteristics of financial reporting quality (e.g., the neutrality element of reliability).[10]

---

[8] For example, former SEC Chairman Levitt defined financial reporting quality only in broad terms as a "commitment to integrity and transparency in the way...companies report their financial performance" (Levitt 1999, 1).

[9] A review of audit-firm prepared guidance designed for their clients' audit committees and guidance developed for a corporate-governance related web site shows that terminology from SFAC No. 2 (e.g., reliability) often is used in describing issues that might be discussed in a financial reporting quality meeting with auditors. Thus, it is reasonable to expect that audit committee members would be familiar with the terms of relevance, reliability, and consistency, but not necessarily familiar with the SFAC No. 2 framework. While the participants in this study were not current audit committee members, they were exposed to the terms related to the SFAC No. 2 quality characteristics, although not the actual conceptual framework, in their class study of corporate financial accounting.

We investigate whether experts and literates differ in two judgments related to the quality of the reporting treatment for ten selected financial statement items. First, given audit committee members' responsibilities to be proactive and probing in financial reporting quality discussions with auditors (Turner 2001, 3), we examine whether experts and literates identify different reporting issues for discussion with the auditors. Second, we compare their evaluations of the quality of the reporting treatment of ten selected items with respect to specific elements of reporting quality (e.g., predictive value).

Based on our previous argument, we expect that financial experts' experienced-based, well-developed schemas lead them to raise concerns about the reporting treatments of items often associated with the risk of misstatement, and thus poor reporting quality (Palmrose and Scholz 2001; Moriarty and Livingston 2001). This expectation is consistent with research documenting that individuals with significant financial experience have knowledge of misstatement frequencies for financial statement accounts (Ashton 1991), as well as the misstatements' underlying causes (Libby 1985; Solomon et al. 1999). In addition, archival evidence shows that misstatements are likely to occur in accounts associated with routine business activities such as sales and accounts receivable at least as often as in accounts associated with nonrecurring business activities such as restructurings (Kruetzfeldt and Wallace 1986; Wright and Ashton 1989; Kinney and Martin 1994; KPMG 2000; Nelson et al. 2002; Moriarty and Livingston 2001; Palmrose and Scholz 2001; Turner 2001). This research suggests that when identifying reporting concerns and evaluating the reporting treatment of specific items with respect to reporting quality, experts will be more likely than literates to focus on items associated with recurring activities.

In contrast, because they lack experienced-based schemas, we expect that literates' identifications and evaluations of reporting issues will be influenced more by the salience of the financial items from one of two sources. The first source relates to the premise that literates will obtain episodic knowledge (and potentially develop schemas) about reporting issues predominantly from second-hand sources (e.g., the business press). The acquisition of such knowledge likely is influenced by the public prominence of the issues. Specifically, psychology research indicates that individuals' frequency assessments are formed using the availability heuristic, in which instances that are easier to recall are perceived as more frequent (Tversky and Kahneman 1973). The ease with which individuals recall items can be influenced by factors unrelated to their true frequency, including their vividness and the degree to which issues related to these items are publicized (Lichtenstein et al. 1978). Thus, literates will emphasize financial items prominent in the business press such as restructuring charges.

The second source of salience in identifying reporting issues is based on the proposition that literates will rely more on attributes of the presented information itself and their attention will be drawn to distinctive or unusual items. Research indicates that individuals tend to notice unexpected (Payne et al. 1993, 180–181) or distinctive (Taylor 1982, 190) information, for example, reporting of infrequent activities such as reorganizations and litigation settlements. Infrequent activities also may be emphasized by their financial statement presentation, e.g., income from discontinued operations or the effect of a change in an accounting principle (Maines and McDaniel 2000), and discussions of such items in footnotes and the MD&A (e.g., accounting changes). Thus, literates will be more likely to attend to financial items related to nonrecurring activities in their identifications and evaluations related to financial reporting quality.

---

[10] SFAC No. 2 identifies elements of reporting quality that underlie the three qualitative characteristics. Relevance includes two elements, predictive value for the future and feedback value for past performance; reliability includes elements of verifiability, neutrality, and representational faithfulness (which encompasses comprehensiveness of disclosure). Hereafter, we refer to these elements of reporting-quality characteristics or elements associated with financial reporting quality simply as "elements" or "elements of reporting quality."

In summary, we posit that experts and literates will differ in their identifications of critical issues for discussion with the auditors and in their evaluations of the quality of the reporting treatments for specific financial items. Financial literates, with less experience with specific reporting issues, will be more likely than experts to identify critical reporting issues based on salience. Specifically, we predict that literates will be more likely than experts to identify issues that are both prominent and nonrecurring. The experienced-based episodic knowledge and better-developed schemas of experts will cause them to be less influenced by salience than literates, and thus to select recurring activities of low prominence more often than literates.

This leads to the following alternative-form hypothesis, stated as a joint hypothesis comparing experts' and literates' judgments since our theoretical development is based on relative, rather than absolute, behavior. Additionally, in an experimental setting with restrictions on the number of provided financial items, there are limits on the number of reporting issues participants can list, which results in an interdependence between issues raised and not raised. For example, if participants are limited to identifying a single reporting issue, the decision to discuss a high prominence, nonrecurring item, by design, precludes raising a low prominence, recurring item.

> **H2:** In identifying critical financial reporting issues from a set of financial statements for discussion with the external auditors, financial literates will be more likely than financial experts to identify reporting issues about activities that are prominent in the business press and nonrecurring, while financial experts will be more likely than financial literates to identify reporting issues about recurring activities of low prominence.

Similarly, in evaluating the quality of the reporting treatment for specific items presented in the financial statements with respect to selected elements of reporting quality, we predict that literates will rate the quality of the elements (e.g., predictive value) lower than experts for high prominence, nonrecurring financial items. Conversely, we predict experts' quality ratings of these elements will be lower than literates for low prominence, recurring items.

> **H3a:** In evaluating the quality of the reporting treatment for specific items presented in the financial statements, financial literates will rate the elements of reporting quality lower than financial experts for items related to high prominence, nonrecurring activities.

> **H3b:** In evaluating the quality of the reporting treatment for specific items presented in the financial statements, financial experts will rate the elements of reporting quality lower than financial literates for items related to low prominence, recurring activities.

III. EXPERIMENTAL METHOD

The study employs a 2 (financial designation) × 3 (financial item salience classification) mixed design, with financial designation as a between-subjects factor and financial item salience classification as a within-subjects factor. The between-subjects factor represents two levels of financial designation: financial expertise and financial literacy. Our objective was to select two groups of participants with significant and similar amounts of overall business-related experience, one with experience in accounting/finance and the other with experience in other functional areas or general business experience. The first group, referred to as "financial experts," includes 20 practicing audit managers from a Big 5 CPA firm, who averaged 8.33 years of public accounting experience and 0.50 years of general business experience. This group is a surrogate for the audit committee member who is a financial expert under the Blue Ribbon Committee's and the exchanges' guidelines. The other group, referred to as "financial literates," comprises 18 recent Executive M.B.A. program graduates who had taken at least one financial and one managerial accounting course, which provided the requisite understanding of the three basic financial statements. The Executive M.B.A.s averaged 9.31 years of general business experience outside of accounting and 0.31 years of experience in accounting or finance. This group is a surrogate for the literate but not expert audit committee member.

Participants received financial statements and related information associated with an audit committee review of quarterly (10-Q) financial disclosures prior to their release. Each was instructed to assume the role of an audit committee member for a mid-cap NYSE textile company. The task was to review the first quarter financial statements, MD&A, and disclosure information in preparation for a discussion with the independent auditor about the 10-Q report. As part of this task, participants read background information on the textile industry and information related to the fictional company, including a description of its products, factors affecting revenues, and factors affecting costs.

The financial statements contained ten specific items for which participants subsequently evaluated the quality of the item's reporting treatment with respect to elements of reporting quality. In selecting and creating descriptions of these items, we consulted both a controller of a mid-cap textile manufacturer and an audit partner for a mid-cap textile company engagement, and examined the financial statements of several textile companies. Based on these discussions and our analysis, we selected ten items that are both likely to affect reporting quality for textile companies and classifiable along two salience dimensions (prominence/recurring). We classified each item into one of three salience combinations that represents the within-subjects' manipulation: high prominence/nonrecurring, high prominence/recurring, and low prominence/recurring.[11] Table 1 summarizes the ten items.[12]

The first category (high prominence/nonrecurring) includes a realignment initiative, an

---

[11] We did not include any low prominence/nonrecurring financial statement items. None of the financial items initially identified by the controller and partner or subsequently identified from our examinations of other firms' financial statements in this industry fell naturally within this classification. Nonrecurring items have been extensively covered in SEC policy initiatives, the POB Report (1994), and/or by the Blue Ribbon Committee (1999), with the result that these items are not realistically viewed as having low prominence.

[12] To assess the prominence of the ten financial statement items, we did two things. First, we read transcripts of all 41 speeches given by the SEC chairman, commissioners, or staff during the first five months of 2001, searching for these ten specific items. In addition, we conducted keyword searches for these ten items in the context of financial reporting quality in the *Wall Street Journal, Business Week, Fortune*, and *Barron's* for the period January 2000 to April 2001. The search results indicate that all items designated as high prominence were mentioned at least twice as many times as any item designated as low prominence. Second, we asked 11 auditors (as experts) and five first-year E.M.B.A. students (as literates) who did not participate in the study to classify each of the ten items as to its prominence and recurring nature on scales from 1 (low) to 7 (high). Based on the average assessments, both groups classified realignment, litigation settlement, and revenue recognition issues as the most prominent. The literates rated the accounting reclassification as the next highest in terms of prominence, while the experts' average assessment for reclassification was preceded by the availability of debt financing. Each group classified the realignment, reclassification, and litigation settlement as the least recurring.

TABLE 1
Descriptions of the Selected Financial Statement Items and their Potential Reporting-Quality Implications Classified by Salience Dimension[a]

Reporting Item	*Description of Financial Item and Potential Reporting-Quality Implications*
	High Prominence/Nonrecurring
Realignment initiative	The Company implemented an efficiency initiative two years ago to improve its competitiveness in light of foreign competition. The initiative involved a reduction of sales, manufacturing, and administrative staffs. The cost was reflected in income two years ago, with most of the implementation "essentially complete" by the 4th quarter of the prior year. Of the $20 million reserve established two years ago, a $1.5 million balance remains at the current quarter-end. *This remaining reserve might be questioned as being an excess reserve based on the completion status of the initiative.*
Reclassification of SGA expense	In the current quarter, the Company changed the criteria for items to be included in selling, general, and administrative (SGA) to conform to industry practice. This change resulted in a reclassification of about $3.0 million from SGA to Cost of Sales in the current quarter. The company restated prior year income statements to reflect the change. *The justification for and the timing of the accounting change might be questioned in a discussion of reporting quality.*
Settlement of class action litigation	In the prior year, a group of employees had filed a lawsuit against the company in connection with their terminations due to the realignment initiative. Although management believes that the Company would have prevailed in court, it reached a settlement with these employees to avoid the cost of protracted litigation. The settlement amount of $0.9 million is reflected in current quarter earnings. *A concern about the likelihood of other related litigation and the need to disclose that risk might be raised in a discussion of reporting quality.*
	High Prominence/Recurring
Revenue alliance with Laksmi	The Company has an alliance with Laksmi, an Indian company, in which Laksmi purchases denim from the Company. In return, the Company markets sportswear produced by Laksmi outside the Indian sub-continent. The final price paid by Laksmi for the denim varies inversely with the level of sales over a 12-month period. Revenues are recorded at estimated prices, with adjustments made for changes in estimates. *The reasonableness of these estimates could be a topic for a discussion of reporting quality.*
Revenue related to Mexican joint venture	The Company entered into a joint venture with a production facility in Mexico five years ago. Denim fabric for orders is sent from U.S. plants to the Mexican facility, which assembles the garments and ships orders to the customer. The Company records revenue based on weekly planned shipment schedules, with adjustments for deviations from the schedule made as needed. *The reasonableness of this recording of revenue and the timing of the related adjustments could be raised in a discussion of reporting quality.*

(*Continued on next page*)

TABLE 1 (Continued)

	Low Prominence/Recurring
Renewal of contracts for key customers	Three customers account for approximately 60–65 percent of sales revenue. The Company has three-year purchase agreements with these companies. One of these agreements is up for renewal in the 3rd quarter of the current year. Management believes the relationship has been positive, but there is no guarantee of renewal. *A discussion of reporting quality might include the likelihood of this renewal and the adequacy of management's discussion of the matter in the disclosures.*
Collection of accounts receivable	*Nothing in the financial statements indicated anything unusual with the presentation or valuation of accounts receivable on the current balance sheet or with consistency with prior years. Collectibility might be a concern given the estimates and adjustments related to the revenue recognition items described above, and because the failure to renew the major purchase agreement might also indicate a lower realizable value of the associated receivables.*
Cotton purchase program	The Company has entered into cotton purchasing programs to ensure an uninterrupted supply of cotton. Some contracts leave the Company subject to commodity price fluctuations, with the general trend of increasing cotton prices likely to continue. *The adequacy of the disclosure of this risk could be questioned in a discussion of reporting quality.*
Future debt financing	The information provided in the financial statements showed the balance of long-term and short-term debt, payments due over the next five years, and sources of debt (8.5 percent debentures, term loans, and working capital line of credit). *Debt financing is key to success in this industry, so the ability to obtain future financing and the associated impact on continued viability and profitability could be important in discussions about reporting quality.*
Property, plant, and equipment transactions	*Nothing in the financial statements indicated anything unusual with property, plant, and equipment. Little detail was provided about property, plant, and equipment, however, so in a discussion of reporting, quality questions could be raised about the investment program and the related disclosures.*

---

[a] Financial statement items are listed by Prominence/Recurring category. Participants received background information that included these ten financial items and were asked to assess the quality of the reporting treatment for each of the ten specified items with respect to six elements underlying the quality characteristics. Statements in italics in this table reflect potential reporting-quality implications for each item; all but the italicized statements were included in the experimental materials.

accounting change (expense reclassification), and a litigation settlement related to the realignment. The second category (high prominence/recurring) concerns revenue recognition—a recurring activity emphasized recently by the SEC and covered heavily in the business press. The first item relates to a revenue alliance with a company in India; the second relates to revenues associated with a joint venture in Mexico. Finally, the third category (low prominence/recurring) consists of five recurring business activities that have received relatively little exposure in the press. The activities include a purchase program for raw materials, collection of accounts receivable, a contract renewal for a key customer, future debt financing, and transactions in property, plant, and equipment. We chose reporting treatments that raise quality concerns such as the accuracy of estimates in the financial statements, the reasonableness of an accounting change, and adequacy of disclosure.[13]

After reading the financial statement information, participants answered two sets of questions. The first question set elicited responses on 11-point scales for the overall assessments of financial reporting quality and assessments of qualitative characteristics of relevance, reliability, and comparability for the financial statements taken as a whole. Participants also identified three critical reporting issues to be discussed in the audit committee's meeting with the external auditors and listed their specific concerns related to each issue.[14]

In the second question set, participants assessed the quality of the reporting treatment of each of the ten specific reporting items presented in Table 1 with respect to six elements of reporting quality. The elements were the degree to which the reporting treatment (1) provides knowledge about future operations/economic events, (2) provides knowledge about past operations/economic events, (3) provides for comprehensive and transparent presentation, (4) reflects the company's business activities, (5) is neutral, and (6) is based on measures with little uncertainty and few assumptions. The first two elements underlie the relevance quality characteristic, while the last four are elements of reliability.[15] Each question included a description of the element to ensure a common interpretation and the responses were elicited on seven-point scales from low to high.[16] Finally, participants provided demographic information.

IV. RESULTS

Table 2 presents descriptive data on participants' backgrounds and self-evaluations of their own business and financial expertise. The self-evaluations are consistent with participants' actual experience levels (previously reported) and our designations of experts and literates. As seen in Panel B of Table 2, both groups report similar perceived levels of business expertise (means of 4.44 vs. 4.40 for literates and experts, respectively, on a seven-point scale; $t = 0.15$, $p < .880$), but literates rate their financial expertise significantly lower than the experts (means of 2.64 and 5.15, respectively; $t = -7.38$, $p < .001$). In addition, the financial literates perceive themselves more as general business experts than financial experts ($t = 7.33$, $p < 0.001$ one-tailed), while experts perceive themselves more as financial

---

[13] All financial amounts related to the ten financial statement items exceeded typical quantitative benchmarks for materiality, i.e., 5 percent of pretax income, which was $300 (thousand) for the fictional textile company.

[14] We restricted both the total number of financial items presented to participants and the number of critical reporting issues to be discussed. We made this design choice given realistic constraints on the amount of time participants could spend reading and completing the experiment. The restriction is not unrealistic in that the number of reporting issues available for discussion for companies is relatively unlimited yet the meeting time between the audit committee and auditors for discussion generally is fixed. In addition, the SEC recommends that firm advisors (including audit committees) "should identify the three, four or five most critical accounting [issues]" (Pitt 2001).

[15] We did not ask participants to evaluate the comparability of the ten financial statement items since only one of our specified items, the reclassification due to an accounting change, relates to comparability across years. Additionally, participants did not receive financial statements for comparative companies.

[16] We expected that our participants would be familiar with the definitions of relevance, reliability, and comparability based on their classroom experiences and exposure to the popular press; however, because the specific names labeling the elements underlying the reliability and relevance characteristics are not used consistently, we provided descriptions of the elements within the experimental materials. In our descriptions, we avoided the use of specific SFAC No. 2 terminology that would likely be more familiar to experts and instead used terminology consistent with the elements listed in SFAC No. 2 and Jonas and Blanchet (2000). For example, the description of the "representational faithfulness" component asked participants to assess the degree to which information accurately reflects the company's business activities. See Table 5 for complete descriptions.

TABLE 2

Descriptive Statistics: Participants' Work Experience and Expertise Self-Assessments

Panel A: Experience[a]

	Literates (n = 18)		*Experts (n = 20)*		*Test: Literates vs. Experts*[b]	
	Mean	*Std. Devn.*	*Mean*	*Std. Devn.*	*t*	*Pr > |t|*
Business work experience (years)						
Accounting/Finance	0.31	0.83	8.33	3.61	−9.19	0.001
General business	9.31	5.87	0.50	1.00	6.61	0.001
Total	9.61	5.57	8.83	3.82	0.51	0.612
Number of board positions held	0.17	0.38	0.15	0.37	0.14	0.892

Panel B: Self-Assessments[c]

	Literates (n = 18)		*Experts (n = 20)*		*Test: Literates vs. Experts*[b]	
	Mean	*Std. Devn.*	*Mean*	*Std. Devn.*	*t*	*Pr > |t|*
Business Expert	4.44	0.98	4.40	0.82	0.15	0.880
Financial Expert	2.64	1.16	5.15	0.93	−7.38	0.001
Financially Literate	4.72	1.23	6.25	0.91	−4.39	0.001
Familiar with the textile business	2.17	1.15	2.50	1.79	−0.67	0.505
Expert in Generally Accepted Accounting Principles (GAAP)	2.39	1.38	5.80	0.95	−8.96	0.001

Panel C: Tests of Differences in Self-Assessments within Participant Groups

	t	*Pr > t*[b]
Literates (n = 18)		
Business Expert > Financial Expert	7.33	0.001
Financially Literate > 4 (scale midpoint)	2.50	0.023
4 (scale midpoint) > Financial Expert	4.98	0.001
Experts (n = 20)		
Business Expert < Financial Expert	4.27	0.001
Financially Literate > 4 (scale midpoint)	11.05	0.001
Financial Expert > 4 (scale midpoint)	5.51	0.001

---

[a] "Work experience," measured in years, is reported for (1) full-time accounting-related or finance-related experience, and (2) full-time general business experience that is nonaccounting and nonfinance. "Number of board positions" is the number of Board of Director member positions held currently or in the past.

[b] Test results for testing for differences in mean responses of financially literate individuals and financial experts. Wilcoxon tests of the distributions provide qualitatively equivalent results.

[c] Self-assessments are reported on a seven-point scale (ranging from 1 = completely disagree to 7 = completely agree). Participants were asked the extent to which they agree with each of the statements: "I consider myself [a(n)]..."

than general business experts (t = 4.27, $p < .001$). While literates believe they are financially literate (i.e., the mean of 4.72 is significantly greater than the scale midpoint of 4.0; t = 2.50, $p < .023$), they do not believe they are financial experts (i.e., the mean of 2.64 is significantly less than 4.0; t = 4.98, $p < .001$). In contrast, those designated as experts consider themselves to be financial experts (i.e., the mean of 5.15 is greater than the scale midpoint of 4.0; t = 5.51, $p < .001$) and to be more expert in generally accepted accounting principles than do literates (5.80 vs. 2.39, respectively; t = –8.96, $p < .001$). Finally, the two groups do not differ in their perceived familiarity with the textile industry (2.17 and 2.50 for literates and experts, respectively; t = –0.67, $p < .505$), suggesting that any differences in their assessments of financial reporting quality are unlikely to be due to differences in industry-specific knowledge.[17]

Assessments of Overall Financial Reporting Quality

Hypothesis 1 predicts that financial experts' evaluations of overall financial reporting quality will be more strongly associated with their assessments of relevance, reliability, and comparability relative to the evaluations of financial literates. We first present descriptive data related to participants' overall reporting quality and the three quality characteristics assessments, and then provide specific tests of H1.

Panel A of Table 3 presents participants' assessments of overall financial reporting quality and individual assessments of relevance, reliability, and comparability for the financial statements taken as a whole (on 11-point scales with endpoints labeled "Low" and "High" and the midpoint label "Moderate"). Since we make no predictions about differences between the levels of literates' and experts' quality assessments, we present descriptive results for these variables. Panel A shows that literates assessed overall reporting quality for the financial statements taken as a whole higher relative to experts (8.11 vs. 7.05, respectively; t = 1.99, $p < .054$). While the literates' and experts' assessments of the overall relevance and reliability of the financial reporting do not differ (7.50 vs. 7.55; t = –0.10, $p < .922$ for relevance and 7.39 vs. 6.55; t = 1.37, $p < .178$ for reliability), literates assessed the comparability of the financial reporting lower than the experts (6.00 vs. 7.20, respectively, t = –1.89, $p < .067$). Together, these results suggest that literates' frameworks or schemas for evaluating financial reporting quality may not be internally consistent, or that literates' frameworks include factors not represented by SFAC No. 2 characteristics. First, the literates' mean assessment of overall reporting quality is significantly higher than that of the experts even though literates' assessments for each of the three underlying characteristics are lower than or not statistically different from those of the experts. Second, eight out of 18 literates assessed overall reporting quality higher than any of their three individual characteristics' assessments, while only one out of 20 experts had this same pattern ($p < .007$ based on Fisher's exact test).

To test H1, we regress participants' overall financial reporting quality evaluations on their assessments of relevance, reliability, and comparability to evaluate whether experts' reporting quality evaluations more strongly reflect their assessments of the three underlying quality characteristics than do literates' evaluations. Panel B of Table 3 reports individual regression results for literates and experts, while Panel C shows the combined regression that tests for differences between literates' and experts' regression coefficients.

The adjusted R^2 for the literates' model in Panel B of Table 3 is −0.063 and none of the coefficients achieves significance. In contrast, the adjusted R^2 for the experts' model is

---

[17] Untabulated results indicate literates' and experts' perceived familiarity with the textile industry are similar and both are below the response scale midpoint of 4.0 (t = –6.76, $p < .001$ and t = –3.74, $p < .002$, respectively).

TABLE 3
Assessments of Overall Financial Reporting Quality and Underlying Quality Characteristics

Panel A: Assessments of Overall Financial Reporting Quality and Underlying Quality Characteristics[a]

	Literates ($n = 18$)		*Experts ($n = 20$)*		*Test: Literates vs. Experts*[b]	
	Mean	*Std. Devn.*	*Mean*	*Std. Devn.*	*t*	$Pr > \lvert t \rvert$
Financial Reporting Quality	8.11	1.28	7.05	1.90	1.99	0.054
Relevance	7.50	1.65	7.55	1.47	−0.10	0.922
Reliability	7.39	1.54	6.55	2.14	1.37	0.178
Comparability	6.00	1.85	7.20	2.04	−1.89	0.067

Panel B: Regressions for Literates' and Experts' Assessments of Underlying Quality Characteristics on Overall Financial Reporting Quality

$$\text{Financial Reporting Quality} = a_0 + b_1{*}\text{Relevance} + b_2{*}\text{Reliability} + b_3{*}\text{Comparability}$$

	Coefficient Estimates (t-statistics) [p-values][c]						
	Intercept	*Relevance*	*Reliability*	*Comparability*	*F*	*Prob > F*	*Adj. R^2*
Literates (n = 18)	5.82	0.07	0.08	0.19	0.67	0.587	−0.063
	(2.91)	(0.30)	(0.39)	(0.95)			
	[0.011]	[0.384]	[0.351]	[0.178]			
Experts (n = 20)	−1.76	0.84	0.13	0.23	9.09	0.001	0.561
	(−1.00)	(3.99)	(0.85)	(1.36)			
	[0.331]	[0.001]	[0.205]	[0.096]			

Panel C: Combined Regression of Assessments of Underlying Quality Characteristics on Overall Financial Reporting Quality[d]

$$\begin{aligned}\text{Financial-Reporting Quality} = {} & a_0 + a_1{*}\text{Literate} + b_1{*}\text{Relevance} + b_2{*}\text{Literate}{*}\text{Relevance} \\ & + b_3{*}\text{Reliability} + b_4{*}\text{Literate}{*}\text{Reliability} + b_5{*}\text{Comparability} \\ & + b_6{*}\text{Literate}{*}\text{Comparability}\end{aligned}$$

Coefficient Estimates (t-statistics) [p-values][e]

Intercept	*Literate* Intercept*	*Relevance*	*Literate* Relevance*	*Reliability*	*Literate* Reliability*	*Comparability*	*Literate* Comparability*
−1.76	7.58	0.84	−0.77	0.13	−0.05	0.23	−0.04
(−0.98)	(2.86)	(3.90)	(−2.47)	(0.83)	(−0.17)	(1.34)	(−0.14)
[0.334]	[0.008]	[0.001]	[0.010]	[0.207]	[0.432]	[0.096]	[0.443]

(*Continued on next page*)

TABLE 3 (Continued)

---

[a] Characteristics related to overall financial reporting quality were measured on 11-point scales (1 = Low, 11 = High). The specific questions asked for each item are as follows:

Financial Reporting Quality = The quality of financial reporting is typically determined in relation to the usefulness of the financial information to external users (e.g., creditors and investors). Please rate the overall financial reporting quality of Forrest and Sons, Inc.'s 1st quarter 10-Q information;

Relevance = Please rate the overall relevance of Forrest and Sons, Inc.'s 1st quarter 10-Q information;

Reliability = Please rate the overall reliability of Forrest and Sons, Inc.'s 1st quarter 10-Q information; and

Comparability = Please rate the overall comparability of Forrest and Sons, Inc.'s 1st quarter 10-Q information.

[b] Reported results are for tests of differences of overall mean assessments between financially literate individuals and financial experts. Wilcoxon tests of the distributions provide qualitatively equivalent results.

[c] Reported p-values reflect two-tailed tests for the intercept term and one-tailed tests for the estimated coefficients for Relevance, Reliability, and Comparability.

[d] "Literate" is a dummy variable with Literate = 0 for experts and Literate = 1 for literates.

[e] Reported p-values reflect two-tailed tests for the intercept and Intercept*Literate coefficients and one-tailed tests for the estimated coefficients for Relevance, Literate*Relevance, Reliability, Literate*Reliability, Comparability, and Literate*Comparability.

0.561, indicating that experts' assessments of reporting quality with respect to some underlying quality characteristics are aggregated into their overall assessments. Specifically, the coefficient on relevance is significantly positive ($t = 3.99$, $p < .001$, one-tailed) and the comparability coefficient is marginally significant ($t = 1.36$, $p < .096$, one tailed). The coefficient on reliability is not significant at conventional levels ($t = 0.85$, $p < .205$, one-tailed).[18] The combined regression in Panel C shows that the literates' model statistically differs from that of the experts in terms of both the intercept (literate*intercept coefficient = 7.58, $p < .008$) and relevance (literate* relevance coefficient = –0.77, $p < .010$). However, experts' and literates' reliability and comparability coefficients are not statistically different. Overall, these results suggest differences between experts' and literates' financial reporting quality frameworks; specifically, literates' overall reporting-quality evaluations do not aggregate their assessments of the SFAC No. 2 characteristics, while at least some of the qualitative characteristics in the SFAC No. 2 framework are reflected in the experts' overall reporting-quality models.

Sensitivity Analyses

Because of small sample sizes, we examined the regressions used to test H1 for the presence of influential observations using procedures in Belsley et al. (1980), and re-estimated the regressions excluding two influential observations from each group. The adjusted R^2 of the literates' regression increased to 0.224, and the coefficient on relevance became marginally significant ($p < .094$). The adjusted R^2 of the experts' regression increased to 0.670, and all three coefficients were at least marginally significant ($p < .001$ for relevance, $p < .089$ for reliability, and $p < .015$ for comparability). Thus, our general conclusions about the differences between literates' and experts' frameworks for evaluating financial reporting quality are robust to the elimination of influential observations.

---

[18] We also performed regressions including interactions among relevance, reliability, and comparability. None of the interactions was significant for either the literates' or experts' model, consistent with research showing that individuals' judgments typically are based on a linear combination of information (Dawes 1979).

A possible explanation for the literates' failure to incorporate quality characteristics in their overall assessments of financial reporting quality is that they lack knowledge about the SFAC No. 2 definitions of relevance, reliability, and comparability. As indicated earlier, because these participants were directly exposed to, and were expected to be familiar with, these general terms we did not test our participants' knowledge of these definitions. We can provide, however, some evidence of the effects of having a reminder of the definitions of two of the characteristics, relevance and reliability, based on results from a pilot test conducted prior to our experiment. In the pilot test, definitions of relevance and reliability were provided to the 22 participants when we asked for assessments of overall financial reporting quality, relevance, and reliability.[19] Results indicated that the literates' reporting-quality model that included only variables for relevance and reliability exhibited a stronger association with SFAC No. 2 characteristics. Specifically, the coefficients on both relevance and reliability were significantly positive (0.51, $p < .001$ and 0.31, $p < .003$, respectively).[20] However, comparing these results with those of the experts' model based on data from the actual experiment that include only relevance and reliability as independent variables shows the literates' relevance coefficient still is smaller than the comparable coefficient for experts ($p < .049$), and there is no statistical difference between literates' and experts' reliability coefficients ($p < .603$).[21] These results suggest that while providing SFAC No. 2 definitions for relevance and reliability improves the aggregation of literates' assessments of these two quality characteristics into their overall reporting-quality evaluations, literates' models still exhibit differences from experts' models with respect to the weight placed on relevance.

Identification of Critical Reporting Issues and Evaluations of the Elements of Reporting Quality for the Reporting Treatment of Specific Financial Items

Hypothesis 2 predicts that in identifying critical reporting issues for discussion, financial literates will be more likely to select financial items that are prominent in the business press and distinctive (i.e., nonrecurring), while financial experts will be more likely to identify low prominence reporting items related to recurring activities. Similarly, H3a states that in evaluating the quality of the reporting treatment for selected items, literates will provide lower assessments of the elements of reporting quality for the more salient financial items than will experts; H3b predicts that experts will provide lower assessments of elements underlying the quality of the reporting treatment for low prominence, recurring reporting items than will literates. Collectively, these hypotheses predict that we should observe the greatest differences between experts' and literates' identifications and evaluations of specific reporting treatments in the high prominence/nonrecurring vs. low prominence/recurring category (i.e., reporting treatments for items of high salience vs. low salience on both dimensions).

---

[19] The assessment of comparability was excluded as the research question, design, and context of the pilot experiment differed slightly from the experiment used in the primary study. Thus, our pilot results cannot address the effects of providing a definition for comparability.

[20] For the literate participants in the actual experiment, the estimated coefficients with only relevance and reliability included in the model are 0.18 and 0.06, respectively, and neither is significantly different from zero.

[21] For the expert participants in the actual experiment, the estimated coefficients with only relevance and reliability included in the model are 0.93 and 0.22, respectively, which are significantly greater than zero at $p < .001$ and $p < .062$.

Identification of Critical Financial Reporting Issues

For H2, participants identified, in order of importance, three reporting issues they deemed most critical to address in a one-hour reporting-quality discussion with the auditors, and stated the single most important question related to each issue. This exercise was a free-generation task in which participants could list any reporting issue contained in the financial statements, as well as any issue not explicitly presented (e.g., integrity of management). Table 4 presents the percentages of participants in each group who selected a specific reporting issue as their first, second, or third choice for discussion, and the overall percentages for whether an issue was listed among the participants' three choices. Panel A reports participants' responses that correspond to any of the ten specified items identified as having greatest implication for reporting quality by our analysis of textile companies' reports and our discussions with two independent experts, a controller and an audit partner (see Section III). The identified issues in this panel are classified within one of the three prominence/recurring categories. Panel B of Table 4 reports issues participants raised that did not relate to the ten specified items—i.e., issues with lesser reporting-quality implications.[22]

While we predict experts' and literates' identifications will differ most between the high prominence/nonrecurring and low prominence/recurring categories, we first compare the percentage of experts vs. literates who identified concerns related to one of the ten specific financial items, i.e., the percentage of issues raised across the three prominence/recurring categories combined (Panel A of Table 4 subtotal), relative to all other issues identified (i.e., Panel B subtotal). Overall, 81.7 percent (of the combined top three choices) of the experts' selected issues related to the ten specific items, compared to 65.0 percent of the literates. In comparison, 18.3 percent of the experts' combined top issues related to items not within the set of the ten specified items, compared to 35.4 percent of the literates ($p < .034$ based on Fisher's exact test). This same pattern also holds for the identifications of the 1st, 2nd or 3rd issues individually (although differences are not statistically significant).

As shown in Panel B of Table 4, inventory concerns dominated the literates' list of issues unrelated to the ten specific items; primarily, literates raised concerns about the company's use of both the FIFO and LIFO methods of inventory accounting. We did not expect inventory to be a significant issue; thus, literates' identification of inventory as a concern is consistent with H2 in that literates likely perceive the use of both FIFO and LIFO as unusual or distinctive based on their prior exposure to the topic from introductory accounting textbooks that tend to focus on consistency issues and the use of one method or the other for a single company. These results indicate that experts appeared more likely than literates to identify reporting concerns in areas previously identified as important to the reporting quality of textile companies. In contrast, literates were distracted by items not identified as critical to reporting quality.

Next we examine differences between literates' and experts' overall distributions for identifications that relate only to the ten specific reporting items. We focus on participants' first selected reporting issue, since limited time for most discussions between the audit committee and auditors suggests that audit committee members' first questions will receive the most attention (Olson 1999, 6). We also report results for the comparisons of the distributions for all three issues combined for just this subset of identified issues.

---

[22] Two independent coders classified the reporting issues identified by participants into one of the four categories listed in Table 4, based on both participants' indicated issue and specific question. There was 77 percent agreement between the two coders. All disagreements were resolved through discussions.

TABLE 4
Reporting Issues Identified by Participants as Most Critical for Discussion with External Auditors[a]

	Literates (n = 18)				Experts (n = 20)			
	(a)	(b)	(c)	(d)	(a)	(b)	(c)	(d)
	% Responses Listing Issue:				% Responses Listing Issue:			
Identified Reporting Issue	1st	2nd	3rd	Within Top 3	1st	2nd	3rd	Within Top 3
Panel A: Issues Participants Raised that Were Included in the Set of Ten Specified Critical Reporting Issues, by Prominence/Recurring Category[b]								
High Prominence/Nonrecurring								
Realignment initiative	11.1	5.6	11.1	9.3	15.0	15.0	30.0	20.0
SGA reclassification	27.8	5.6	11.1	14.8	—	15.0	10.0	8.3
Settlement of litigation	—	11.1	5.6	5.6	—	5.0	5.0	3.3
Category total	38.9	22.3	27.8	29.7	15.0	35.0	45.0	31.7
High Prominence/Recurring								
Revenue recognition								
Laksmi	16.7	5.6	5.6	9.3	10.0	5.0	—	5.0
Mexican joint venture	—	11.1	—	3.7	5.0	5.0	5.0	5.0
Laksmi & Mexican JV	—	11.1	5.6	5.6	25.0	10.0	5.0	13.3
Category total	16.7	27.8	11.2	18.6	40.0	20.0	10.0	23.3
Low Prominence/Recurring								
Key customer contracts	5.6	5.6	—	3.7	25.0	5.0	15.0	15.0
A/R collection	—	—	—	—	—	—	—	—
Raw materials (cotton)	5.6	5.6	5.6	5.6	—	15.0	5.0	6.7
Debt	—	5.6	5.6	3.7	—	5.0	5.0	3.3
Property, plant & equipment	—	5.6	5.6	3.7	5.0	—	—	1.7
Category total	11.2	22.4	16.8	16.7	30.0	25.0	25.0	26.7
Subtotal—Panel A issues raised	**66.8**	**72.5**	**55.8**	**65.0**	**85.0**	**80.0**	**80.0**	**81.7**

(Continued on next page)

TABLE 4 (Continued)

	Literates (n = 18)				*Experts (n = 20)*			
	(a)	*(b)*	*(c)*	*(d)*	*(a)*	*(b)*	*(c)*	*(d)*
	% Responses Listing Issue:				*% Responses Listing Issue:*			
Identified Reporting Issue	*1st*	*2nd*	*3rd*	*Within Top 3*	*1st*	*2nd*	*3rd*	*Within Top 3*
Panel B: Issues Participants Raised That Were Not Included in the Set of Ten Specified Critical Reporting Issues[c]								
Revenue recognition—vague	—	—	5.6	1.9	—	—	—	—
Risk factors—SFAS No. 133	—	—	5.6	1.9	—	5.0	10.0	5.0
Consolidation	5.6	—	5.6	3.7	—	—	5.0	1.7
Inventory	22.2	16.7	11.1	16.7	—	5.0	5.0	3.3
Statement of cash flows	—	—	5.6	1.9	5.0	—	—	1.7
Estimates and reserves	—	—	—	—	5.0	—	—	1.7
General disclosure detail	—	5.6	5.6	3.7	—	10.0	—	3.3
Management quality	—	—	—	—	5.0	—	—	1.7
Long-term business outlook	5.6	5.6	—	3.7	—	—	—	—
International expansion	—	—	5.6	1.9	—	—	—	—
Subtotal—Panel B issues raised	33.4	27.9	44.7	35.4	15.0	20.0	20.0	18.3
Total (Panels A and B combined)[d]	100	100	100	100	100	100	100	100

---

[a] We instructed participants that the audit committee and the external auditor planned to meet for approximately one hour to discuss issues with the greatest potential influence on the company's financial reporting quality for the quarter's 10-Q. We asked participants to list their three most critical issues (in order of priority) to discuss with the external auditor, and for each issue, to specify the single most important question they wished to address. Independent coders, without knowledge of the participant's group (i.e., literate vs. expert), categorized all responses. There was 77 percent agreement between the two coders; all disagreements were resolved through discussions.

[b] Panel A includes responses corresponding to the ten specific items identified as having greatest implications for financial reporting quality by our analysis of textile companies' reports and our discussions with two independent experts, a controller and an audit partner. Responses are classified within one of three prominence/recurring categories.

[c] Panel B includes responses that did not correspond to the ten specific items identified as having greatest implications for financial reporting quality by our analysis of textile companies' reports and our discussions with two independent experts.

[d] Columns may not sum to 100 due to rounding.

Columns (a), (b), and (c) report the percentage of each participant group that listed the specific issue as the 1st, 2nd, or 3rd topic, respectively, for discussion with the external auditor; (d) reports the percentage of each participant group's responses that listed the issue as any of the top three choices.

A comparison of Columns (a) in Panel A of Table 4 supports the general expectation that the frequency of issues selected first across the three salience categories differs between literates and experts. Specifically, 38.9 percent of the literates' first issues fall within the high prominence/nonrecurring category compared to 15.0 percent for the experts, while

16.7 percent (40.0 percent) of the literates (experts) raised concerns in the high prominence/recurring category as their first issue. In addition 11.2 percent of the literates' compared to 30.0 percent of the experts' primary concerns related to low prominence/recurring business activities. The distribution of frequencies for these three categories differs at the .10 level between experts and literates (based on Fisher's exact test).

Hypothesis 2 predicts that literates, relative to experts, will be more likely to identify as critical those issues that are salient on both dimensions (i.e., high prominence/nonrecurring), and experts will be more likely than literates to identify concerns related to low prominence/recurring items. The results in Panel A of Table 4 support this prediction for the first-raised critical reporting issue. Specifically, 38.9 percent (15.0 percent) of literates (experts) raised concerns related to a high prominence/nonrecurring item as their first selection while 11.2 percent (30.0 percent) of literates (experts) identified a low prominence/recurring item as their first choice. A test of literates' and experts' first issues within these two categories shows that the two groups have different patterns across the two categories ($p < .077$ based on Fisher's exact test).[23] These results suggest that experts and literates focus on different types of reporting issues in identifying their primary financial reporting quality concerns to discuss with the external auditor.

As previously noted, the distributions considering all three issues identified by participants (including issues with lesser reporting-quality implications) differed between experts and literates. However, excluding the issues raised that are not included among the ten specified items, the distribution of frequencies across all salience categories for all three issues identified (i.e., Column (d) for literates and experts in Panel A of Table 4) does not differ between literates and experts ($p < .755$ based on Fisher's exact test). A test of literates' vs. experts' combined three choices within the high prominence/nonrecurring and low prominence/recurring categories does not support H2 ($p < .314$ based on Fisher's exact test).[24]

The results based on Table 4 suggest that experts and literates focus on different types of financial reporting issues in evaluating financial reporting quality. Overall, experts are more likely to identify issues consistent with the ten financial items identified as important by our analysis of textile companies' reports and our discussion with two experts, while literates raise concerns about items with less impact on reporting quality. Experts and literates also differ with respect to the single issue they consider most critical when considering only the set of ten selected financial reporting items: experts are more likely than literates to identify ongoing activities as their most critical issue, while literates identify prominent, nonrecurring issues. When all three reporting issues listed by participants within this set of ten items are considered, however, there is no difference between groups.

Evaluation of Quality Elements of the Reporting Treatment for Specific Financial Items

Hypotheses 3a and 3b predict that financial literates will rate the quality of elements underlying the reporting treatment for specific financial items lower than experts for items categorized as high prominence/nonrecurring and financial experts will evaluate lower quality than literates for items in the low prominence/recurring category, respectively. Table 5 reports the literates' and experts' assessments on seven-point scales of the two elements underlying the relevance characteristic and four elements of reliability, and the averages for the overall relevance and reliability characteristics for the reporting treatment of the ten specific items grouped according to the item's salience category. To test H3a and H3b, we examine the averages of the participants' ratings for the respective quality elements underlying the characteristics of relevance and reliability, since our theory does not predict which elements will be affected for individual salience categories.

---

[23] A comparison of the identification rate for high prominence/nonrecurring items vs. all others indicates literates are more likely than experts to raise these issues for the first choices listed ($p < .097$, one-tailed, based on Fisher's exact test). A comparison of the identification rate for low prominence/recurring items vs. all other issues indicates experts are not more likely to raise these items than literates ($p < .118$, one-tailed, for the first issue raised based on Fisher's exact test).

[24] A comparison of the identification rate for high prominence/nonrecurring items vs. all others indicates literates are not more likely than experts to raise these issues among the three issues raised ($p < .669$ based on Fisher's exact test), nor are experts more likely than literates to identify low prominence/recurring items vs. all others ($p < .144$ for all three issues raised based on Fisher's exact test).

Hypothesis 3a predicts that literates, relative to experts, will assign lower quality ratings to the reporting treatment for high prominence/nonrecurring items (all p-values in Panel A of Table 5 are one-tailed given the directional prediction of H3b). As reported in Panel A of Table 5, literates' average relevance assessment for the reporting treatments of these items was lower than that of experts (4.66 vs. 5.09; $t = -1.72$, $p < .047$). Literates' average assessment over all four reliability elements also was lower than that of the experts (4.15 vs. 4.61; $t = -1.67$, $p < .052$). Finally, relative to experts, literates had lower overall combined ratings for the elements underlying the relevance and reliability characteristics for the reporting treatment of the high prominence/nonrecurring items (4.32 vs. 4.77, respectively; $t = -2.09$, $p < .022$).

Results in Panel C of Table 5 are consistent with the prediction that experts will ascribe lower quality to reporting treatments of the low prominence/recurring items than will literates (all p-values are one-tailed given the directional prediction of H3b). The average of experts' assessments for the two relevance elements is significantly below that of the literates (3.42 vs. 4.07, respectively $t = -2.45$, $p < .001$). Experts also assessed lower values for the reliability elements for these items, as indicated by their lower average of 3.67 vs. 3.95 for the literates ($t = -1.54$, $p < .067$). Finally, experts' average evaluation of all six relevance and reliability elements combined for the low prominence/recurring category is lower than that of literates (3.55 vs. 3.97; $t = -2.31$, $p < .014$).

Our theory provides no specific prediction for the relation between experts' and literates' assessments for the high prominence/recurring items (i.e., revenue recognition items). Experts should raise concerns about these items because of their association with misstatements, while literates should be concerned because of these items' prominence. Results in Panel B of Table 5 suggest there are no statistically reliable differences between experts' and literates' assessments of the quality of the reporting treatments except for the neutrality element, where literates ascribed greater neutrality than experts to the reporting treatment of the items ($p = .042$). The average assessment for relevance and reliability combined for experts does not differ relative to that of literates (3.78 vs. 3.25, respectively; $t = 1.54$, $p < .134$); neither do the separate average assessments for relevance (3.87 vs. 3.50, respectively; $t = 0.92$, $p < .363$) or reliability (3.70 vs. 3.11, respectively; $t = 1.62$, $p < .114$).

Overall, these results suggest that differences between literates' and experts' knowledge not only direct their attention to different reporting issues, but also affect their assessments of the quality of the reporting treatment of specific financial items. Consistent with our expectations, literates (relative to experts) assign a lower quality rating to the reporting treatment of items that have received prominence in the business press and are unusual in their nature or presentation, and experts provide lower quality ratings than literates for recurring activities with low prominence.

TABLE 5
Participants' Mean Assessments for Elements of Reporting-Quality Characteristics for Specific Financial Items by Recurring/Prominence Categories[a]

Panel A: Mean Assessments for High Prominence/Nonrecurring Financial Items

	Mean Assessment			
Component[b]	*Literates*	*Experts*	*t-value*[c]	*Pr < t*
Relevance components				
Future	4.51	4.47	0.11	0.545
Past	4.80	5.72	−3.04	0.002
Relevance average	4.66	5.09	−1.72	0.047
Reliability components				
Comprehensive	4.43	5.32	−2.77	0.004
Reflects business activities	4.08	4.32	−0.54	0.295
Neutral	3.35	4.11	−1.58	0.061
Little uncertainty	4.72	4.75	−0.08	0.468
Reliability average	4.15	4.61	−1.67	0.052
Overall average	4.32	4.77	−2.09	0.022

Panel B: Mean Component Assessments for High Prominence/Recurring Financial Items

	Mean Assessment			
Component	*Literates*	*Experts*	*t-value*	*Pr > |t|*
Relevance components				
Future	3.50	3.30	0.42	0.678
Past	4.25	3.70	1.02	0.316
Relevance average	3.87	3.50	0.92	0.363
Reliability components				
Comprehensive	3.28	3.23	0.11	0.913
Reflects business activities	4.65	3.75	1.68	0.102
Neutral	3.89	3.00	2.11	0.042
Little uncertainty	2.89	2.63	0.54	0.591
Reliability average	3.70	3.11	1.62	0.114
Overall average	3.78	3.25	1.54	0.134

Panel C: Mean Component Assessments for Low Prominence/Recurring Financial Items

	Mean Assessment			
Element	*Literates*	*Experts*	*t-value*	*Pr > t*
Relevance elements				
Future	4.06	3.58	1.44	0.079
Past	4.20	3.33	3.06	0.002
Relevance average	4.07	3.42	2.45	0.001
Reliability elements				
Comprehensive	3.51	3.10	1.29	0.103
Reflects business activities	4.04	3.75	1.05	0.151
Neutral	4.13	3.93	0.70	0.245
Little uncertainty	3.93	3.78	0.94	0.177
Reliability average	3.95	3.67	1.54	0.067
Overall average	3.97	3.55	2.31	0.014

(Continued on next page)

TABLE 5 (Continued)

---

[a] For each of the ten financial items described in Table 1, participants separately rated each of the six elements of reporting-quality characteristics. For example, for "Future," participants were instructed to rate as "7" the economic issue for which the financial information provides the highest relative degree of knowledge about future operations and/or future economic events. Then they were to rate as "1" the item for which the financial information provides the lowest relative degree of knowledge about future operations and/or future economic events. Finally, they were asked to rate each of the remaining items between 7 (high degree) and 1 (low degree), inclusive, relative to the other nine financial reporting items. Precise wording provided to participants for each of the six elements evaluated is as follows:

Future = the degree to which financial information provides knowledge about future operations and/or future economic events;
Past = the degree to which financial information provides knowledge about past operations and/or past economic events;
Comprehensive = the degree to which financial information is presented in a comprehensive and transparent manner;
Reflects business activities = the degree to which financial information appropriately reflects the company's business activities (e.g., operating or investing activities);
Neutral = the degree to which financial information is presented by management in a neutral manner; and
Little uncertainty = the degree to which financial information is measured with little uncertainty and is based on few assumptions.

[b] "Relevance" is the average of "Future" and "Past" assessments, and "Reliability" is the average of "Comprehensive," "Reflects business activities," "Neutral," and "Little uncertainty." "All elements" is the average of all six individual element assessments.

[c] Reported t-value results are for tests of differences between overall mean assessments of Literates vs. Experts. In Panel A (C) the prediction is that mean assessments for Literates (Experts) will be lower than for Experts (Literates), so the p-value is the probability of a t-statistic that is more negative (positive) than the observed value. There is no directional prediction for Panel B comparisons, so we report two-tailed p-values.

V. DISCUSSION AND CONCLUSION

This paper explores some possible effects of rules related to audit committee members' qualifications and offers evidence on the question of how individuals "assess the quality of a company's reporting" (Jonas and Blanchet 2000, 358). Specifically, we provide evidence on how financial expertise vs. financial literacy affects individuals' overall assessments of reporting quality, their identifications of critical reporting issues that warrant discussion with auditors, and their assessments of the quality of the reporting treatment for specific financial statement items.

Our results suggest experts' frameworks or schemas of financial reporting quality reflect some of the quality characteristics (e.g., relevance) espoused by the FASB's SFAC No. 2. On the other hand, financially literate individuals' assessments of overall financial reporting quality do not reflect these characteristics consistently, even though these individuals were exposed previously to the SFAC No. 2 characteristics. The finding that experts incorporate relevance into their overall financial reporting quality assessments to a greater degree than literates is important in that information relevance is at the heart of several recent SEC and FASB initiatives. For example, fair value reporting and forward-looking information are both considered ways to improve the relevance of information for investors even if reliability is reduced. Given that audit committees are required to evaluate and discuss the overall quality of a company's financial reporting, the results suggest that inclusion of financial experts on audit committees is likely to add structure to the discussion of overall reporting quality and improve the consistency of assessments of overall reporting quality.

Experts and literates also differ in their identifications of reporting issues for discussion with the auditors and their assessments of the quality of selected reporting treatments, suggesting that experts and literates bring different perspectives to the discussions with auditors about financial reporting quality. Specifically, financial experts appear to identify as most critical for discussion those issues that are related more to recurring activities, i.e., activities which their experiences would suggest are associated with quality concerns (e.g., restatements). In contrast, financial literates initially select issues for discussion based on items that have received prominent coverage in the business press and are nonrecurring/ distinctive in nature or have less important implications for reporting quality. While both types of items are important for the overall quality of financial reporting and for overall business risk, our results suggest that including experts on the audit committee is likely to bring a focus on issues that may be accorded lower priority by audit committee members who are only financially literate.

This paper raises several issues for future research. First, auditors served as financial experts in our experiment. While auditors meet the exchanges' definition of a financial expert, it is not clear that other types of financial experts (e.g., chief financial officers) would make the same reporting-quality judgments as did the auditors in our study. There is no such concept as a "generic" financial expert; each type of expert brings specific experiences to the task of assessing financial reporting quality. Further research could investigate whether and how various types of financial experts differ in their reporting-quality judgments.

Recent survey evidence indicates that auditors believe audit committee member diversity in terms of both financial and nonfinancial experiences leads to more effective discussions about reporting quality (Krishnamoorthy et al. 2002). Our findings suggest that experts and literates bring at least somewhat differing perspectives to discussions and evaluations of financial reporting quality, but our study does not investigate how other types of experience and knowledge (e.g., general business, industry) affect audit committee effectiveness. Future research could examine whether and how such nonfinancial experience might provide benefits to audit committee performance.

Future research also could examine whether the results of this study extend to audit committees in an interactive group setting. Although our study did not address the effects related to possible group dynamics among experts and literates within an audit committee setting, we argue that our results related to differences in individuals have implications for audit committees for several reasons. First, while groups tend to outperform individuals, the best individual typically will outperform the group as a whole (Hastie 1986). Additionally, research suggests that the accuracy of an interacting group's judgment is a function of the group's ability to recognize the most accurate individual in the group and weight that person's judgments more heavily (Libby et al. 1987; Sniezek 1989). Given the financial expert in an audit committee will be designated as such *a priori* (rather than the audit committee assessing this during the group's interaction), the expert's evaluation of reporting quality likely will be weighted more heavily in the audit committee's discussion than the judgments of the financially literate audit committee members. Thus, understanding individual experts' evaluations of reporting quality is relevant in understanding how the expertise requirement might affect the audit committee's discussion or judgments.

Future research could use interacting groups to investigate directly whether the results of our study have implications for audit committees, specifically whether the inclusion of a financial expert leads to a focus on reporting-quality issues that are not addressed by a committee comprised only of literates. Additionally, because group performance depends on many factors (Tindale 1989), it is possible that the selection of issues for the financial reporting quality discussion may be less diverse if literates defer to the judgment of the expert, rather than raising their own issues. Thus, future research could include a more in-depth investigation of the effects of expertise in an interactive audit committee group setting.

REFERENCES

American Institute of Certified Public Accountants (AICPA). 1999. *Audit Committee Communications.* Statement on Auditing Standards No. 90. New York, NY: AICPA.

———. 2000. *Quality of Accounting Principles—Guidance for Discussions with Audit Committees.* Practice Alert 2000-2. New York, NY: AICPA.

Anderson, J. R. 2000. *Cognitive Psychology and Its Implications.* New York, NY: Worth Publishers.

Ashton, A. H. 1991. Experience and error frequency knowledge as potential determinants of audit expertise. *The Accounting Review* 66 (April): 218–239.

Association for Investment Management and Research (AIMR). 1999. Letter from the Financial Accounting Policy Committee to the Blue Ribbon Committee on Improving the Effectiveness of Corporate Audit Committees. January 28, 1999. Charlottesville, VA: AIMR.

Belsley, D. A., E. Kuh, and R. E. Welsch. 1980. *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity.* New York, NY: John Wiley & Sons, Inc.

Blue Ribbon Committee on Improving the Effectiveness of Corporate Audit Committees. 1999. *Report and Recommendations of the Blue Ribbon Committee on Improving the Effectiveness of Corporate Audit Committees.* New York, NY: New York Stock Exchange and The National Association of Securities Dealers.

Braiotta, L., Jr. 1999. *The Audit Committee Handbook.* 3rd edition. New York, NY: John Wiley & Sons, Inc.

Castellan, N. J., Jr. 1993. *Individual and Group Decision Making.* New York, NY: Lawrence Erlbaum Associates.

Costales, S. B., and G. Szurovy. 1994. *The Guide to Understanding Financial Statements.* New York, NY: McGraw-Hill, Inc.

Dawes, R. M. 1979. The robust beauty of improper linear models in decision making. *The American Psychologist* 34 (July): 571–582.

Dechow, P. M., R. G. Sloan, and A. P. Sweeney. 1996. Causes and consequences of earnings manipulation: An analysis of firms subject to enforcement actions by the SEC. *Contemporary Accounting Research* 13 (Spring): 1–36.

DeFond, M. L., and J. Jiambalvo. 1991. Incidence and circumstances of accounting errors. *The Accounting Review* 66 (July): 643–655.

DeZoort, F. T. 1998. An analysis of experience effects on audit committee members' oversight judgments. *Accounting, Organizations and Society* 23 (January): 1–21.

———, and S. Salterio. 2001. The effects of corporate governance experience and financial reporting and audit knowledge on audit committee members' judgments. *Auditing: A Journal of Practice & Theory* 20 (September): 31–47.

Ernst & Young. 2000. *Audit Committees: Implementing the New Rules.* New York, NY: Ernst & Young LLP.

Financial Accounting Standards Board (FASB). 1980. *Qualitative Characteristics of Accounting Information.* Statement of Financial Accounting Concepts No. 2. Stamford, CT: FASB.

Fiske, S. T., and S. E. Taylor. 1991. *Social Cognition.* New York, NY: McGraw-Hill.

Hastie, R. 1986. Review essay: Experimental evidence on group accuracy. In *Information Pooling and Group Decision Making: Proceedings of the Second University of California, Irvine, Conference on Political Economy,* edited by B. Grofman and G. Owens. Greenwich, CT: JAI Press.

Jenkins, E. L. 2000. Letter to the Auditing Standards Board. April 18. Norwalk, CT: FASB.

Jonas, G. J., and J. Blanchet. 2000. Assessing quality of financial reporting. *Accounting Horizons* 14 (September): 353–363.

Kinney, W. R., Jr., and R. D. Martin. 1994. Does auditing reduce bias in financial reporting? A review of audit-related adjustment studies. *Auditing: A Journal of Practice & Theory* 13 (Spring): 149–155.

Kirk, D. J. 2000. Experiences with the Public Oversight Board and corporate audit committees. *Accounting Horizons* 14 (March): 103–111.

KPMG Audit Committee Institute. 2000. *Audit Committee Survey*. New York, NY: KPMG LLP.

Krishnamoorthy, G., A. Wright, and J. Cohen. 2002. Audit committee effectiveness and financial reporting quality. *Australian Accounting Review* (forthcoming).

Kruetzfeldt, R. W., and W. A. Wallace. 1986. Error characteristics in audit populations: Their profile and relationship to environmental factors. *Auditing: A Journal of Practice & Theory* 6 (Fall): 20–43.

Levitt, A. 1998. The numbers game. Speech to the NYU Center for Law and Business, New York, NY, September 18.

———. 1999. Quality information: The lifeblood of our markets. Speech to The Economic Club of New York, New York, NY, October 18.

———. 2000. Speech before the Conference on the Rise and Effectiveness of New Corporate Governance Standards, New York, NY, December 12.

Libby, R. 1985. Availability and the generation of hypotheses in analytical review. *Journal of Accounting Research* 23 (Autumn): 648–667.

———, K. T. Trotman, and I. Zimmer. 1987. Member variation, recognition of expertise, and group performance. *Journal of Applied Psychology* 72 (1): 81–87.

———, and J. Luft. 1993. Determinants of judgment performance in accounting settings: Ability, knowledge, motivation, and environment. *Accounting, Organizations and Society* 18 (5): 425–450.

———. 1995. The role of knowledge and memory in audit judgment. In *Judgment and Decision-Making Research in Accounting and Auditing*, edited by R. H. Ashton, and A. H. Ashton, 176–206. Cambridge, U.K.: Cambridge University Press.

Lichtenstein, S., P. Slovic, B. Fischhoff, M. Layman, and B. Combs. 1978. Judged frequency of lethal events. *Journal of Experimental Psychology: Human Learning and Memory* 4 (November): 551–578.

Maines, L. A., and L. S. McDaniel. 2000. Effects of comprehensive-income characteristics on nonprofessional investors' judgments: The role of financial-statement presentation format. *The Accounting Review* 75 (April): 179–207.

McMullen, D. A. 1996. Audit committee performance: An investigation of the consequences associated with audit committees. *Auditing: A Journal of Practice & Theory* 15 (Spring): 87–103.

Moriarty, G. B., and P. B. Livingston. 2001. Quantitative measures of the quality of financial reporting. *Financial Executive* 17 (July/August): 53–56.

Nelson, M. W., J. A. Elliott, and R. L. Tarpley. 2002. Evidence from auditors about managers' and auditors' earnings-management decisions. *The Accounting Review* (Supplement): 175–202.

Olson, J. F. 1999. How to make audit committees more effective. *The Business Lawyer* 54 (May): 1097–1113.

Palmrose, Z-V., and S. Scholz. 2001. The accounting causes and legal consequences of non-GAAP reporting: Evidence from restatements. Working paper, University of Southern California.

Paskell-Mede, M., and P. Jackson. 1999. Meeting expectations. *CA Magazine* 132 (November): 49–50.

Payne, J. W., J. R. Bettman, and E. J. Johnson. 1993. *The Adaptive Decision Maker*. Cambridge, U.K.: Cambridge University Press.

Pitt, H. L. 2001. How to prevent future Enrons. *Wall Street Journal* (December 11): A18.

PricewaterhouseCoopers LLP. 2000a. *Audit Committee Update 2000*. New York, NY: PricewaterhouseCoopers LLP.

———. 2000b. *Audit Committee Effectiveness—What Works Best*. 2nd edition. Altamonte Springs, FL: The Institute of Internal Auditors Research Foundation.

Public Oversight Board (POB) Advisory Panel on Auditor Independence. 1994. *Strengthening the Professionalism of the Independent Auditor*. Stamford, CT: POB.

Securities and Exchange Commission (SEC). 1940. *In the Matter of McKesson & Robbins, Inc.* Accounting Series Release No. 19. Washington, D.C.: Government Printing Office.

———. 1999. *Audit Committee Disclosure*. Release No. 34–42266. December 22. Washington, D.C.: Government Printing Office.

Smith, E. R. 1998. Mental representations and memory. In *The Handbook of Social Psychology*, 4th edition, edited by D. T. Gilbert, S. T. Fiske, and G. Lindzey, 391–440. Boston, MA: McGraw-Hill.

Sniezek, J. A. 1989. A examination of group process in judgmental forecasting. *International Journal of Forecasting* 5: 171–178.

Solomon, I., M. D. Shields, and O. R. Whittington. 1999. What do industry-specialist auditors know? *Journal of Accounting Research* 37 (Spring): 191–208.

Taylor, S. E. 1982. The availability bias in social perception and interaction. In *Judgment under Uncertainty: Heuristics and Biases*, edited by D. Kahneman, P. Slovic, and A. Tversky, 190–200. Cambridge, U.K.: Cambridge University Press.

Tindale, R. S. 1989. Group vs. individual information processing: The effects of outcome feedback on decision making. *Organizational Behavior and Human Decision Processes* 44: 454–473.

Turner, L. E. 2001. Quality, transparency, accountability. Speech to the Financial Executives Institute-AICPA's "Benchmarking the Quality of Earnings" Conference, New York, NY, April 26.

Tversky, A., and D. Kahneman. 1973. Availability: A heuristic for judging frequency and probability. *Cognitive Psychology* 4 (September): 207–232.

Wild, J. J. 1996. The audit committee and earnings quality. *Journal of Accounting, Auditing and Finance* 11 (Spring): 247–276.

Wright, A., and R. H. Ashton. 1989. Identifying audit adjustments with attention-directing procedures. *The Accounting Review* 64 (October): 710–728.

Yates, J. F., 1990. *Judgment and Decision Making*. Englewood Cliffs, NJ: Prentice Hall, Inc.

Zacharias, C. A. N. 2000. New rules, new responsibilities. *Journal of Accountancy* 190 (August): 53–55.

THE ACCOUNTING REVIEW
Vol. 77
Supplement 2002
pp. 169–173

DISCUSSION OF Evaluating Financial Reporting Quality: The Effects of Financial Expertise vs. Financial Literacy

Mark E. Peecher
University of Illinois at Urbana–Champaign

I. INTRODUCTION

McDaniel et al. (2002) provide an experimental examination of financial reporting quality judgments that contrasts the perspectives of two types of participants, namely audit managers and recent E.M.B.A. graduates. They argue that these two types of participants reasonably proxy for audit committee members with financial expertise and financial literacy, respectively. Their experimental examination demonstrates some ways in which *semantic* and *episodic* knowledge differences likely influence the perspectives taken when people complete ill-structured tasks involving financial reporting quality assessments.[1] Their examination also is timely, as considerable diversity of opinion exists among corporate governance commentators and regulators about the net benefits of requiring audit committees to include financial literates and at least one financial expert. By providing theory-based empirical evidence about systematic financial reporting quality judgment differences between financial experts and financial literates, their paper may prove useful to both academics and corporate governance commentators and regulators.

Specifically, McDaniel et al. experimentally test three hypotheses that predict financial experts and literates to hold different, albeit complementary, perspectives on financial reporting quality. Hypothesis 1 predicts that, relative to financial literates, financial experts rely more systematically on judged relevance, reliability, and comparability when judging overall reporting quality. Hypotheses 2 and 3 predict that financial experts rely less on issues recently covered by the business media, both when judging the relative criticality of financial reporting issues to discuss with auditors (H2) and when evaluating dimensions of relevance and reliability for specific accounting issues (H3).

---

[1] Within permanent memory, semantic knowledge is stored as generalized abstractions independent of the circumstances surrounding its acquisition, e.g., time, place, what one was doing, whereas the storage of episodic knowledge includes such circumstances (Tulving 1972). With repeated instances of similar judgments, the human mind abstracts relevant portions of episodic memory into accessible semantic form to reduce cognitive effort associated with subsequent judgments.

I appreciate the helpful comments provided by Steve Salterio on this commentary.

The theoretical premise for all three hypotheses is that financial experts have relatively well-developed memories containing semantic (and episodic) knowledge that they can readily access to render judgments about reporting quality, whereas financial literates have relatively impoverished memories containing mostly episodic knowledge. Applying this premise to H1, McDaniel et al. conjecture that financial experts possess greater working knowledge of the conceptual framework (SFAC No. 2) than do financial literates. For H2 and H3, they conjecture that financial experts make reporting-quality judgments based on a memory that is more comprehensive, with respect to financial reporting quality, than that possessed by financial literates. Specifically, financial experts' memories reflect both coverage by the business press and their own rich sources of knowledge. In contrast, and largely by default, financial literates' memories impound information from the business press, which seems to emphasize nonrecurring items.

Their paper reports three main findings. One, compared to financial literates, financial experts rely more systematically on judged relevance (but not on judged reliability and judged comparability) when assessing overall financial reporting quality. Two, when judging the criticality of reporting issues, financial literates attend to issues that are prominent (in terms of business press coverage) and nonrecurring more than do financial experts. Three, when judging subcomponents of reporting quality related to relevance and reliability, financial literates make systematically lower judgments for accounting issues that are prominent/nonrecurring than do financial experts.

Conference attendees raised questions about alternative ways to interpret some of the experimental findings and about whether the paper's theory is sufficiently rich to generalize to real-world audit committees. In this discussion, I integrate questions raised by the Conference attendees with my own views. I begin with some general comments that apply to all three hypotheses and end with comments directed to specific hypotheses.

II. GENERAL COMMENTS

While Conference attendees did not dispute that audit managers and graduates of an Executive M.B.A. program meet NYSE (and other) requirements for designation as financial experts and financial literates, respectively, they voiced concern about how well these two "convenience" samples represent real-world audit committees. On the surface, this concern calls for research on the relative frequency with which such persons actually populate audit committees.[2] Beneath the surface, this concern calls for a careful assessment of differences between the participants in this study's two convenience samples and alternative participants who would also qualify as financial experts (and financial literates). In the absence of theoretically important differences, there would be no compelling reason beyond scientific replication to recruit actual audit committee members as participants.

One potentially important theoretical difference is that some real-world audit committee members, especially (former) audit managers, may possess greater industry expertise than that possessed by the experimental participants. Some Conference attendees wondered whether it would be beneficial to measure the influence of financial literacy/expertise, after controlling for a range of industry literacy/expertise. Others asked whether we could be sure that any participants attained a level of performance that could reasonably be construed as expert performance on the experimental tasks.

---

[2] Recent archival empirical research suggests that, at least historically, audit committee members typically have been neither audit managers nor E.M.B.A. students and that less than 25 percent of them are CPAs (Beasley and Salterio 2001). Thus, if, going forward, audit managers (or other CPAs) increasingly serve on audit committees, then this would be a significant change in the composition of audit committees.

While such issues are tangential to McDaniel et al.'s research question, they raise questions about whether audit committee members should *uniformly possess* industry knowledge or expertise (in addition to financial literacy or expertise). Indeed, in recent years large audit firms have modified their professional development and audit approaches to ensure that industry expertise and financial expertise are cultivated *jointly* in auditors. Thus, it is reasonable to ask whether auditors who lack industry knowledge would be willing to be designated as an audit committee financial expert.

Further, as Conference attendees noted, industry literacy/expertise may partly compensate for shortcomings in financial sophistication, at least in terms of developing questions to pose to auditors about activities that are peculiar to the industry. Whether such audit committee members would be capable of understanding auditors' responses in a financial accounting sense is more uncertain (Salterio and DeZoort 2001).

It is thus plausible that important boundary conditions—in particular, industry knowledge—at least sometimes prevail in real-world audit committees that would reduce or nullify the reported inter-participant type differences observed in the reported experiment. Future experimentation could examine the net benefits of joint attributes such as financial literacy and industry knowledge, regardless of the requirements for audit committee membership. By documenting differences related to reporting-quality assessments across financial experts and financial literates, when both groups presumably lack industry expertise, the current study sets the stage for the next steps of scholarly inquiry.

III. COMMENTS ON HYPOTHESIS 1

Some Conference attendees expressed a belief that there was little *a priori* uncertainty as to whether judged relevance, reliability, and comparability would explain more variation in the overall financial reporting quality assessments of financial experts than in the assessments of financial literates. Other Conference attendees noted that one way to strengthen the test of H1 would have been to ask participants to specify a number of qualitative attributes that they believe strongly influence financial reporting quality. This approach may have revealed a need to move beyond the attributes discussed in SFAC No. 2 to address both reporting quality and differences across participant types that were masked by the explicit provision of relevance, reliability and comparability.

IV. COMMENTS ON HYPOTHESES 2 AND 3

Conference attendees questioned the extent to which the paper's discussion of episodic and semantic memory offers better *a priori* support for predictions of *intra-*, as opposed to *inter-*, *participant type* differences. As an example, if financial literates' memory-based schemas were based largely on coverage in the business press, then an *intra-participant type* expectation would be warranted. Financial literates likely identify more nonrecurring issues that have been extensively and prominently covered by the business press than recurring or nonrecurring issues that have not received such coverage.[3] Financial experts, in contrast, likely would have abstracted semantic knowledge from multiple sources, so it seems less likely that coverage by the business press would have much influence on the issues they identify.

---

[3] This assumes that the business media does, in fact, stress nonrecurring issues more than recurring issues and that participants' recollections of their experiences with the business media are representative of actual media coverage. Since the instrument did not ask participants to recall the nature of accounting issues covered in the news, there is some slippage in the mapping from actual media coverage to the relative salience of issues (i.e., salience is a concept that applies to participants' minds).

Interestingly, though, regardless of whether financial experts are influenced by business press coverage, they could well identify a greater number of nonrecurring items than could financial literates, simply because the experts have more knowledge and can therefore identify more issues (in absolute numbers) of all types. Literates, in contrast, likely can identify a greater ratio of nonrecurring issues that have received press coverage relative to issues that have not. Consequently, as some Conference attendees noted, the theory advanced by McDaniel et al. arguably provides only ambiguous support for predicting an *inter-participant type* difference such that financial literates are likely to identify more nonrecurring issues than are financial experts.

Other Conference attendees expressed concern over difficulties in assessing the benefits and costs of financial literates placing a larger emphasis on nonrecurring items relative to recurring items. While McDaniel et al. note that misstatements and restatements likely occur as least as often in the reporting of recurring business activities as in the reporting of nonrecurring business activities, the conditional probability of misstatement or restatement may be greater *given* a nonrecurring (i.e., less frequent) activity as opposed to a recurring (i.e., more frequent) business activity. As in general diagnostic-inference settings, an unusual symptom (e.g., a nonrecurring business activity) likely should trigger a revision of unconditioned, base-rate beliefs about possible causes (e.g., misstatements) and effects (e.g., restatements). Regardless of the business press focus on certain nonrecurring activities, the conditional probability of restatement given nonrecurring activities may rise above the analogous conditional probability given recurring activities if the nonrecurring activities are more unusual.

V. CONCLUSION

McDaniel et al. (2002) provide experimental evidence about how requiring audit committees to include both financial literates and at least one financial expert could influence the perspectives taken in completing ill-structured financial reporting quality assessments. Evidence about how, and how well, audit committee surrogates complete these tasks within a controlled setting may help auditors, audit committees, investors, and regulators assess the net benefits of requiring both financial literacy and expertise of audit committee members and of stipulating what constitutes acceptable evidence of financial literacy/expertise.

Future researchers can build on this timely study in several ways. For example, future research could use recall tests to investigate the degree to which memory differences likely drive the experimental findings reported in this study, recruit industry experts to test whether industry literacy/expertise interacts with financial literacy/expertise to influence performance on tasks involving financial reporting quality, and mimic audit committee interactions by teaming up at least one financial expert with one or more financial literates.

REFERENCES

Beasley, M., and S. Salterio. 2001. The relationship between board characteristics and voluntary improvements in audit committee composition and experience. *Contemporary Accounting Research* 18 (Winter): 539–570.

McDaniel, L. S., R. D. Martin, and L. A. Maines. 2002. Evaluating financial reporting quality: The effects of financial expertise vs. financial literacy. *The Accounting Review* (Supplement): 139–167.

Salterio, S., and T. DeZoort. 2001. The effects of corporate governance experience and financial reporting and audit knowledge on audit committee members' judgments. *Auditing: A Journal of Practice & Theory* 20 (September): 31–47.

Tulving, E. 1972. Episodic and semantic memory. In *Organization and Memory*, edited by E. Tulving and W. Donaldson. New York, NY: Academic Press.

THE ACCOUNTING REVIEW
Vol. 77
Supplement 2002
pp. 175–202

Evidence from Auditors about Managers' and Auditors' Earnings Management Decisions

Mark W. Nelson
John A. Elliott
Cornell University
Robin L. Tarpley
The George Washington University

ABSTRACT: This paper reports analyses of data obtained using a field-based questionnaire in which 253 auditors from one Big 5 firm recalled and described 515 specific experiences they had with clients who they believe were attempting to manage earnings. This approach enables us to analyze separately managers' decisions about how to attempt earnings management and auditors' decisions about whether to prevent earnings management by requiring adjustment of the financial statements. Our results indicate that managers are more likely to attempt earnings management, and auditors are less likely to adjust earnings management attempts, which are structured (not structured) with respect to precise (imprecise) standards. We also find that managers are more likely to make attempts that increase current-year income, but auditors are more likely to require that those attempts be adjusted, that managers are more likely to make attempts that decrease current-year income with unstructured transactions and/or when standards are imprecise, and that auditors are more likely to require adjustment of attempts that they identify as material or that are attempted by small clients.

We appreciate comments from Michael Gibbins (the discussant), Katherine Schipper (the editor), and Linda Bamber, Mike Bamber, Rob Bloomfield, Karl Hackenbrack, Frank Hodge, Steve Kachelmeier, Bill Kinney, Lisa Koonce, Susan Krische, Bob Libby, John Lyon, Steve Salterio, two anonymous reviewers, and participants at *The Accounting Review*'s Conference on Quality of Earnings and the Conference on Behavioral Research in Financial Reporting (both at Emory University), as well as workshop participants at the Australian Graduate School of Management, University of Connecticut, Cornell University, McMaster University, Michigan State University, Northwestern University, Oxford University, SUNY at Buffalo, Temple University, The University of Texas at Austin, and Washington University. We also appreciate research assistance provided by Jennifer Bremner, Shana Clor, Cathy Nelson, and Steve Smith, financial support provided by Cornell's Johnson Graduate School of Management, and the access to audit partners and managers provided by an anonymous Big 5 firm. A prior version of this paper was titled, "Where Do Companies Attempt Earnings Management, and When Do Auditors Prevent It?"

Keywords: *earnings management; auditing; financial accounting standards; precision; rule-based standards; transaction structuring; materiality; independence.*

Data Availability: *Contact the authors.*

I. INTRODUCTION

Earnings management can be defined as non-neutral financial reporting in which managers intervene intentionally in the financial reporting process to produce some private gain (Schipper 1989). Managers can intervene by modifying how they interpret financial accounting standards and accounting data, or by timing or structuring transactions (Healy and Wahlen 1999).[1] Because many such interventions are difficult to distinguish from appropriate applications of GAAP, the definition of earnings management hinges fundamentally on managerial intent, which is difficult to assess using *ex post* accounting information (Dechow and Skinner 2000). As a consequence, researchers have struggled to identify earnings management, and accounting research has not provided much evidence about the characteristics of accounting standards that encourage earnings management, or about the extent to which various aspects of the financial reporting process discourage earnings management (Healy and Wahlen 1999).

This paper reports analyses of data obtained using a field-based questionnaire in which 253 audit partners and managers from one Big 5 firm recalled and described 515 specific experiences they had with clients who they believed were attempting to manage earnings. Auditors are required by GAAS to understand their clients' incentives and to search for differences between actual and expected performance that may indicate misstatements, so the auditors who participated in our study were relatively well positioned to identify specific instances of earnings management. Because respondents provided transaction-level data about attempts covering a range of financial accounting transactions, including attempts that are purely judgmental as well as attempts that involve transaction structuring, these data allow us to examine how attempts are affected by the precision of financial accounting standards[2] and by other characteristics of attempts. Unlike studies that focus on only post-audit information, we consider separately managers' decisions about how to attempt earnings management and auditors' decisions about whether to require adjustments.[3] Thus, our study provides evidence about how a key feature of accounting standards (precision of rules) and a key feature of the financial reporting process (activity of external auditors) influence earnings management.

Results of descriptive analyses indicate that the earnings management attempts in our sample occurred in numerous accounting areas, including revenue recognition, business combinations, intangibles, fixed assets, investments and leases, but by far the most frequently identified attempts involve reserves. Respondents believe that managers' attempts were motivated by a variety of incentives, including the need to meet analysts' estimates and influence the stock market, to reach targets set by compensation contracts or debt covenants, to communicate economic information to stakeholders, and to smooth income or improve future income, as well as by combinations of these incentives. Auditors adjusted

---

[1] We use "transaction structuring" to include management attempts to manage earnings by modifying contracts, transactions, or activities.

[2] We use the term "precise standard" to refer to a criterion that specifically allows or disallows a particular accounting treatment or that is specified numerically.

[3] We use "adjustment" to denote an auditor-required accounting treatment or a refusal to issue an unqualified opinion.

44 percent of the attempts in our sample, 21 percent were not adjusted because the auditor believed the client demonstrated compliance with GAAP, 17 percent because the auditor did not have convincing evidence the client's position was incorrect, and the remaining 18 percent because of some other reason (usually, immateriality).

Consistent with theory and results from law, economics, and accounting, our hypothesis tests provide evidence that managers were more likely to attempt earnings management (and auditors were less likely to adjust attempts) that were either structured to comply with precise standards or unstructured with respect to imprecise standards. Consistent with prior research in accounting, managers were more likely to attempt earnings management that increased current income, and auditors were more likely to adjust these attempts. To the extent that managers made attempts that *decreased* current income, they did so with respect to imprecise standards and/or with unstructured transactions, presumably to retain maximum flexibility for making future income-increasing attempts that reverse the current-period decrease in income. Consistent with other research in accounting, auditors were more likely to adjust attempts that were more material and/or that were attempted by smaller clients.

We believe these results are useful for understanding how earnings management attempts of various types move through the financial reporting process. For example, attempts involving reserves tend to be unstructured, governed by judgmental standards, and income decreasing, so they are less likely to be adjusted. Attempts involving leases, consolidations and determining the appropriateness of the equity vs. cost method tend to be governed by precise standards, and transactions in these areas are amenable to structuring, so they also are less likely to be adjusted. In contrast, although attempts involving revenue recognition occur relatively frequently in our sample, they tend to increase current income, and they are more likely to be adjusted.

Overall, we identify several factors that affect how managers attempt earnings management and whether auditors adjust attempts. Understanding these factors can help participants in the financial reporting process predict where attempts and adjustments occur, and to predict the effects of changes in regulations or audit methodologies that modify these factors. Our results also provide a baseline for assessing the effectiveness of recent regulatory changes (e.g., SAB Nos. 99–102, SAS Nos. 89 and 90) that occurred after we collected data.

The rest of this paper proceeds as follows. In Section II we describe how our research approach complements other approaches, and describe theory and results from law, economics, and accounting that support our hypotheses about factors that affect where earnings management is attempted and whether auditors adjust attempts. In Section III we describe our method. In Section IV we describe analyses and results. In Section V we summarize our findings, discuss implications and limitations, and suggest directions for future research.

II. BACKGROUND AND HYPOTHESES

Contrasting Research Approaches

Prior researchers have used multiple approaches to examine earnings management. For example, "unexpected accruals" studies infer earnings management if total accruals or some subset differs from expectations in the direction favored by an incentive proxy (e.g., Jones 1991; Dechow et al. 1995; McNichols and Wilson 1988). "Distributional" studies test whether the distribution of earnings around benchmarks differs in some predicted way from what would be expected in the absence of earnings management (e.g., Hayn 1995; Burgstahler and Dichev 1997; Burgstahler and Eames 1998; Degeorge et al. 1999; Beatty et al. 2002). "Enforcement-release" studies examine SEC Accounting and Auditing Enforcement Releases (e.g., Dechow et al. 1996; Beneish 1997; Bonner et al. 1998; Panel on

Audit Effectiveness 2000). "Adjusting-entry" studies examine auditors' decisions with respect to proposed adjusting journal entries, some of which may involve earnings management (e.g., Wright and Wright 1997; Kinney and Martin 1994). "Experimental" studies examine earnings-management-related decisions made in laboratory settings (e.g., Hackenbrack and Nelson 1996; King 2002).

While prior research has provided evidence on managers' incentives for earnings management, there is relatively little evidence on the characteristics of accounting standards or transactions that influence where and how managers attempt to manage earnings (Healy and Wahlen 1999). Most unexpected-accruals studies examine aggregate accruals, and distributional studies focus on reported earnings, so they cannot examine particular transactions. Some unexpected-accruals studies and experimental studies focus on particular accruals, but their results may not generalize beyond the particular transaction, account, or industry examined (McNichols 2000; Libby et al. 2001). Enforcement-release studies highlight only the cases the SEC chooses to pursue aggressively, so likewise may not generalize to other types of earnings management. Adjusting-entry studies focus on individual transactions, but include items unrelated to earnings management (e.g., unintentional errors) and exclude important types of earnings management (e.g., successful transaction structuring).

In addition, prior research has struggled to examine how managers' and auditors' decisions with respect to earnings management combine and/or offset in the financial reporting process. Analysis of post-audit financial statements, a joint product of managers and auditors, cannot separate managers' and auditors' decisions. To the extent researchers are interested in the link between managers' incentives and their actions, a focus on post-audit data introduces noise from auditors' adjustment decisions. To the extent researchers are interested in whether auditors mitigate earnings management (e.g., Francis and Krishnan 1999; Bradshaw et al. 2001), a focus on post-audit data understates audit effectiveness by ignoring attempts that auditors detect and prevent. Experimental studies have primarily focused on auditors' adjustment decisions (Libby et al. 2001), and so provide little evidence about managers' attempts. Enforcement-release studies omit instances where auditors prevented earnings management, and both enforcement-release and adjusting-entry studies omit important types of earnings management.

We use a field-based questionnaire to elicit auditors' recollections of specific incidents where managers attempted to manage earnings (see Gibbins and Newton 1994; Gibbins et al. 2001 for other examples of the use of field-based questionnaires). This approach provides transaction-level data about earnings management attempts and auditors' adjustment decisions. Therefore, we can make progress in disentangling managers' and auditors' decisions by examining hypothesized relations involving managers' decisions about how to attempt earnings management (conditional on the decision to make an attempt that the auditor detected) and auditors' decisions about whether to adjust an attempt (conditional on identification of the attempt).

A field-based questionnaire relies on respondents providing accurate recollections. Numerous analyses of eyewitness testimony and other recalled experiences highlight the fallibility of human memory (Loftus 1982), but also indicate circumstances in which recall is relatively accurate (Davies 1993). Techniques have been developed for eliciting accurate recollections, (e.g., Flanagan 1954; Gibbins 2001). For example, accurate recall is encouraged by eliciting simple facts concerning recent events, as opposed to asking leading questions or eliciting elaborate details about long-past events. Intentional inaccuracy is discouraged by highlighting the importance of the research and assuring anonymity. As described in more detail later in the paper, we follow these and other procedures to minimize noise or bias due to relying on recalled data, and also consider the potential for noise or bias when interpreting our results.

Hypotheses

The second and fourth columns of Table 1 summarize our hypotheses. (The results discussed in Section IV are also summarized in Table 1).

Precision of Standards and Transaction Structuring

Research in law and economics examines how rule characteristics affect behavior (see Kaplow [2000] for a review). This literature emphasizes how rule precision affects the degree to which people consider themselves vulnerable to "error costs" (e.g., fines, damages, imprisonment) that result from being found to have violated a rule. When applying *imprecise* rules, people assess high exposure to error costs because adjudication is unpredictable (Shavell 1987), and they may react to this uncertainty by being more careful to comply with the spirit of the rule (Kessler and McClellan 1996). People often lack the expertise necessary to interpret the complexity that typically accompanies precise rules, and so they may act as if they are applying an imprecise rule (Diver 1983; Kaplow 2000). Alternatively, people dealing with precise rules may incur costs to reduce exposure to error costs (Ogus 1992). For example, they might hire experts to interpret the rules (Ehrlich and Posner 1974), or they might incur costs to structure arrangements to evade legal norms (Kaplow 2000). Such investments in experts and/or arrangement structuring are less likely

TABLE 1
Analyses of Factors Affecting Managers' and Auditors' Earnings Management Decisions
Hypotheses and Results

	Sign of Effect on Frequency of Managers' Attempts		*Sign of Effect on Probability Auditor Requires Adjustment*	
Variable	*Hypothesis*	*Result*	*Hypothesis*	*Result*
Precision (P) × Structuring (S) Interaction	H1: +	+**	H2: −	−**
Current-Period Income Effect (CIE)	H3: +	+**	H6: +	+**
P × CIE Interaction	H4: +	+**		
S × CIE Interaction	H5: +	+**		
Materiality			H7: +	+**
Client Size			H8: −	−**

** Significant at $p < 0.05$.

CIE = 1 if current-year-income increasing, −1 if current-year-income decreasing;
Precision = 1 if standard precise, 0 if standard not precise;
Structuring = 1 if transaction structured, 0 if transaction not structured;
Materiality = 1 if material in current or future year individually or in combination with another issue, 0 if never material; and
Client Size = 1 if annual sales < \$50M, 2 if \$50M < sales < \$100M, 3 if \$100M < sales < \$250M, 4 if \$250M < sales < \$1B, 5 if sales > \$1B.

when rules are imprecise, because they are less likely to reduce exposure to error costs significantly (Gifford 1971).

We extend this theory from law and economics to the financial reporting setting, where managers make earnings management attempts and auditors decide whether attempts must be adjusted for the financial statements to receive an unqualified audit opinion. As in the broader legal arena, accounting policymakers must determine the appropriate level of precision and complexity to include in the rules they promulgate (Beresford 1999; Mason and Gibbins 1991), so they must consider how varying levels of precision will affect behavior. We believe managers respond to varying levels of rule precision by varying their earnings management attempts to minimize the chance their auditors will require adjustments.

When accounting rules use precise guidance to distinguish between alternative accounting treatments, and/or to tie accounting recognition to the occurrence of specific events, we predict managers will structure transactions to meet the precise guidance, thereby reducing the risk the auditor will disagree and require adjustments. For example, managers can structure transactions by including elements that meet numerical thresholds; e.g., purchasing "7d insurance" to circumvent the "90 percent of fair-market value" test in SFAS No. 13, para. 7d (Pulliam 1988; Miller and Berton 1993). Managers can also structure transactions by timing events, e.g., recognizing accumulated unrealized gains on "available-for-sale" securities in a particular accounting period by selling the securities as required by SFAS No. 115.[4]

Transaction structuring often requires transaction fees, fees paid to experts for their advice, and/or modification of operational decisions that presumably were optimized prior to the structuring. Therefore, managers should engage in transaction structuring only when its anticipated financial reporting benefits exceed its anticipated out-of-pocket and operational costs. Because the benefits of transaction structuring are less certain. When rules are precise, we predict managers are less likely to structure transactions and more likely to attempt to convince the auditor that they have interpreted the rules correctly. For example, managers may avoid accruing a contingent loss by assigning a relatively high threshold to the term "probable" in SFAS No. 5, and/or interpret the underlying evidence as suggesting a relatively low probability of loss. Managers might state their investment purpose as consistent with the requirements for "available for sale" treatment of investments with respect to SFAS No. 115 to delay income recognition of unrealized gains and losses. Managers applying SFAS No. 87 might assume a relatively low discount rate to minimize pension liability and current-period pension expense.

This extension of theory from law and economics to the accounting context leads us to make the following hypothesis:

> **H1:** Managers are more likely to attempt earnings management with structured transactions when standards are precise and with unstructured transactions when standards are imprecise.

Hypothesis 1 predicates managers' behavior on two assumptions about auditors' behavior, namely, auditors will not require adjustment of structuring that complies with precise rules, and auditors will be willing to discuss, and compromise on, interpretations of imprecise rules. Consistent with this perspective, prior survey evidence indicates that auditors often compromise in client negotiations when accounting standards do not provide precise guidance (Gibbins et al. 2001). In addition, experimental research (in settings where transaction structuring is not possible) indicates that auditors are more likely to permit aggressive accounting positions as the subjectivity of the relevant standards and/or evidence increases.[5] Together with the reasoning supporting H1, these prior results suggest a hypothesis about auditors' decisions that parallels H1:

---

[4] Thus, our definition of transaction structuring includes both "real" earnings management as discussed by Schipper (1989) and "operational" earnings management as discussed by Parfet (2000).

> **H2:** Auditors are less likely to require adjustment of earnings management attempts that are structured when standards are precise and that are not structured when standards are imprecise.

Current-Year Income Effect

Whether a manager makes earnings management attempts that increase or decrease current-year net income presumably depends on his incentives, which generally favor a time series of smoothly increasing income (Barth et al. 1999; DeFond and Jiambalvo 1993). Also, anecdotal evidence suggests that current-income-decreasing earnings management is intended to enable future-income-increasing attempts (e.g., the "cookie-jar reserves" highlighted by Levitt [1998]), but current-income-increasing attempts are typically intended to benefit the current period, rather than enabling future-income-decreasing attempts.[6] Therefore, we predict the following:

> **H3:** Managers are more likely to attempt earnings management that increases current-period income.

Assuming that income-decreasing earnings management is intended to benefit future income (as well as to manage expectations about future performance), it is reasonable to assume that managers desire maximum flexibility for recognizing future income in the periods in which it is needed. Flexibility should be greater under imprecise standards, or when transactions are not structured. Therefore, we predict the following:

> **H4:** Managers are more likely to attempt earnings management that decreases current-period income by using unstructured transactions.

> **H5:** Managers are more likely to attempt earnings management that decreases current-period income when the relevant financial accounting standards are imprecise.

Auditors' incentives include risk of litigation, which research suggest is highest for overstatements of current income and equity (St. Pierre and Anderson 1984; Carcello and Palmrose 1994). Auditing texts encourage auditors to focus on detecting overstatements of current-year net income and equity (Arens and Loebbecke 1997). While audit-adjustment studies indicate that auditors tend to *propose* audit adjustments that reduce income and equity (Wright and Wright 1997; Kinney and Martin 1994), evidence about the adjustments that auditors *require* is less clear. Archival research finds no association between the likelihood of requiring adjustments and the sign of their income effect (Wright and Wright

---

[5] See, e.g., Trompeter (1994), Hackenbrack and Nelson (1996), Kennedy et al. (1997), Nelson and Kinney (1997), Salterio and Koonce (1997), Mayhew et al. (2001), and Kadous et al. (2000).

[6] Earnings management to increase current income by accelerating revenue or postponing expense will decrease future income, but the future decrease is merely a necessary consequence of accrual accounting.

1997), but experimental evidence indicates that auditors are more likely to require adjustment if the unadjusted amount increases income (Braun 2001). Given our understanding of auditors' incentives, we predict the following:

> **H6:** Auditors are more likely to require adjustment of earnings management attempts that increase current-period income.

Materiality and Client Size

Auditors' primary responsibility is to identify and require adjustment of material misstatements, so auditors should be more likely to adjust earnings management attempts as the materiality of those attempts increases (Icerman and Hillison 1991; Wright and Wright 1997; Braun 2001). In addition, qualitatively material attempts might be labeled immaterial to justify not making an adjustment (Libby and Kinney 2000). Therefore, we predict the following:

> **H7:** Auditors are more likely to require adjustment of earnings management attempts they identify as material.[7]

Analytical research theorizes that concern over retaining and attracting clients affects auditors' willingness to require audit adjustments that their clients oppose (for reviews, see Lee and Gu [1998] and Zhang [1999]). This reasoning suggests that, because audit fees increase with client size (Francis and Simon 1987), the probability an adjustment is not made increases with client size, as found in prior audit-adjustment research (Wright and Wright 1997). Also, larger clients are more likely to have the resources to structure transactions carefully and defend aggressive positions effectively. Therefore, we predict the following:

> **H8:** Auditors are more likely to require adjustment of earnings management attempts made by their smaller clients.

III. DATA COLLECTION

Participants

We limited participation to audit partners and managers to ensure that participants had sufficient experience to identify earnings management attempts. Data collection occurred in autumn of 1998.[8] The firm providing access to auditors did not permit us to contact audit managers directly, but did permit us to ask each audit partner we contacted to distribute two additional instruments to managers of their choice. After pretesting the survey with two partners and two senior managers, packets were mailed to 532 audit partners selected randomly from U.S. offices of one Big 5 firm.

---

[7] As discussed further in Section III, we elicited instances of *material* earnings management to insure that our sample includes attempts auditors viewed as important. This does not preclude us from assessing whether auditors' adjustment decisions are explained in part by their *ex post* labeling of attempts as material or immaterial. During the course of negotiations with clients over proposed audit adjustments, auditors could be persuaded by managers that attempts are immaterial, could eventually conclude attempts are immaterial qualitatively or when compared to quantitative thresholds, and/or could label potentially material attempts as immaterial to justify no adjustment. When we base analyses on only the 376 attempts that auditors identified *ex post* as material, results are very similar to those reported.

[8] One month prior to data collection, then SEC Chairman Levitt expressed his concern over earnings management and highlighted specific earnings management techniques. Press coverage of this speech could have affected our subjects' responses. Also, after our data collection, the SEC issued SAB Nos. 99–102, and the ASB issued SAS Nos. 89 and 90, all intended to reduce earnings management. In Section V we discuss the potential effect of these announcements on the generalizability of our results.

Two hundred fifty-three audit partners and managers responded with a total of 525 attempts (10 attempts were dropped due to missing data, so our final sample consists of 515 attempts). Participants had an average of 14.1 years of experience; 43 percent were partners; 57 percent were managers. Our final response rate of 16 percent (20 percent for partners and 14 percent for managers) compares favorably to response rates of mailed surveys of senior financial managers (e.g., 9 percent response rate by Graham and Harvey [2001]; 12 percent response rate by Trahan and Gitman [1995]). However, it is lower than that often achieved in recent mailed survey studies employing auditors as subjects (e.g., 26.7 percent response rate by Gibson and Frakes [1997]; 32 percent response rate by Ayers and Kaplan [1998]).

Several factors probably decreased our response rate. First, asking about sensitive issues generally reduces response rates (Randall and Gibson 1990) and auditors are likely to be sensitive about earnings management.[9] We assured confidentiality and anonymity, but our need to obtain explicit descriptions of actual attempts precluded the use of randomized response techniques (see Warner 1965; Greenberg et al. 1971) that might have further decreased participants' concerns about responding to our survey. Second, our response rate would be reduced if some partners did not distribute surveys to managers or did not encourage managers to respond. This explanation is consistent with our response rate for partners exceeding that for managers. Third, as discussed later in the paper, most auditors were directed to focus on attempts with respect to specific lists of topics to ensure comprehensive coverage of accounting topics. Auditors may not have responded if they had no experiences with respect to those topics, which is consistent with a lower response rate for "list" surveys (15 percent) as compared to "no list" surveys (18 percent) and for managers (who are less experienced than partners). Fourth, auditors with significant not-for-profit experience may not have responded because they believed their experience was not relevant to a study of earnings management, which is consistent with the relatively low amount of experience with not-for-profit clients indicated by our respondents (see footnote 10). Fifth, the survey instrument was lengthy, requiring 30–45 minutes to complete. Responses required significant written elaboration in addition to categorical and numerical answers. A few respondents returned incomplete surveys with a note indicating they lacked sufficient time to finish; others may have simply not responded due to survey length.

One way to assess the potential for nonresponse bias is to compare data from late respondents to data from on-time respondents (Oppenheim 1992; Wallace and Mellor 1988). In our study, 52 "late" responses, received following a reminder, are not significantly different from other responses in any of the analyses reported in the results section. Likewise, there are no significant differences between the first 50 and last 50 responses received, or between the first half and last half of the responses received, providing evidence against nonresponse bias. We also tested for similarity between respondents and the population from which they were drawn (see, e.g., Moore and Reichert 1983). Although confidentiality restrictions prevented us from getting some forms of demographic data, respondents did provide the percentage of their experience that involved clients in each of the many industry classifications used by their firm. The average proportion of auditors' experience in each industry is very similar to the proportion of audit clients the firm has in that industry, suggesting that respondents' backgrounds reflect the general practice of their firm.[10]

---

[9] As an example of sensitivity, one auditor responded with the emphatic statement "I am not now nor have been nor expect to be involved in any way with earnings management."

A final potential concern is that our data collection approach could result in duplicate observations if more than one respondent described the same attempt. Duplication is most likely between partners and managers, as partners presumably would be likely to distribute packets to managers with whom they had worked. To discourage duplication of observations, we ensured that, when auditors were directed to focus on attempts with respect to specific lists of topics (80 percent of our sample), the three surveys mailed to each partner differed in the accounting topics listed. We also reanalyzed all results reported in the paper using only partner data. Results are very similar to those reported.

Initial Instructions and Manipulation of "List" vs. "No List"

Partners received a packet that contained a cover letter from a prominent member of the main office of the firm. The letter included the following statements:

> Researchers from Cornell University are conducting an earnings management study. The project seeks to identify the nature of, and circumstances surrounding, attempts at earnings management. The results of the study will assist our development of an earnings management training tool for less experienced auditors. High-quality input from professionals like you will help to make this study a success.
>
> Please complete the enclosed materials, which ask you to document your experiences.... Please do not delegate this task. Also, please ask two senior managers in your office to complete the packets of materials that are enclosed for them....
>
> No client-specific information will be collected. The data you provide will be kept confidential.

Included in the packet were three envelopes, one addressed "Materials To Be Completed By [partner name]," and two addressed "Materials To Be Completed By A Manager Designated By [partner name]."

The Appendix contains the first page of the survey. We stated that we were interested in learning about participants' experiences with companies that attempt material earnings management. We clarified what we meant by "attempted material earnings management," and then highlighted that earnings management could occur by structuring transactions with respect to standards or by aggressively interpreting standards or data. While all participants were instructed to focus on their most important experiences (i.e., "the most frequently occurring examples of attempted material earnings management" they had encountered), some participants were asked to give experiences with respect to specific accounting topics, and others were not.

As our primary purpose was to consider factors affecting earnings management decisions across a range of transactions, 80 percent of surveys asked participants to consider earnings management with respect to a list of five areas chosen from the 22 areas shown in Table 2. The list was developed in consultation with representatives from the participants' firm to cover the primary non-industry-specific issues dealt with in SFAS Nos. 1–132 and the APBs and ARBs that preceded them. Following the first page of the survey, "list" participants received a table listing five areas and the financial accounting standards related

---

[10] For three industries, there was a statistically significant difference between the mean proportion of respondents' experience and the mean proportion of the firms' clients. Compared to their firms' practice, respondents had more experience with consumer products companies ($p < .001$), and less experience with industrial companies ($p < .001$) and with not-for-profit organizations ($p = .006$). Auditors with predominantly not-for-profit clients may not have responded because they did not believe their experience was relevant to this study.

TABLE 2
Descriptive Data about Earnings Management Attempts

	Number of Earnings Management Attempts (Percent Adjusted by Auditor)					
		Current-Period Income Effect				
Area	*Total*	*Decrease*	*Increase*	*None Identified*	*Percent of Total Governed by Precise Standard*	*Percent of Total that were Structured*
Reserves[a]	178 (35%)	92 (35%)	49 (45%)	37 (24%)	25	0
Revenue Recognition	78 (60)	11 (55)	64 (64)	3 (0)	44	13
Business Combinations	35 (43)	13 (31)	9 (56)	13 (46)	40	14
Non-R&D Intangibles	32 (44)	6 (33)	26 (46)	0 (0)	9	3
Fixed Assets	31 (65)	11 (82)	19 (58)	1 (0)	23	6
Investments	28 (57)	5 (60)	20 (60)	3 (33)	39	14
Leases	23 (39)	2 (0)	18 (39)	3 (67)	83	57
Accounting changes and prior period adjustments	20 (25)	5 (0)	13 (38)	2 (0)	60	5
Compensation	19 (16)	3 (33)	9 (11)	7 (14)	63	26
Consolidations and equity /cost method	15 (27)	1 (0)	12 (33)	2 (0)	47	80
Transfers of Receivables	11 (64)	1 (0)	8 (87)	2 (0)	64	36
Cash Flows	10 (70)	2 (100)	7 (71)	1 (0)	80	10
Taxes	5 (20)	3 (0)	2 (50)	0	60	0
Long-Term Debt	5 (100)	0	4 (100)	1 (100)	80	0
Pensions	5 (20)	3 (33)	2 (0)	0	80	40

(Continued on next page)

TABLE 2 (Continued)

Area	*Number of Earnings Management Attempts (Percent Adjusted by Auditor)*				*Percent of Total Governed by Precise Standard*	*Percent of Total that were Structured*
		Current-Period Income Effect				
	Total	*Decrease*	*Increase*	*None Identified*		
Other Post-Retirement Benefits	4 (25)	0	4 (25)	0	50	0
Segment Reporting	3 (33)	0	0	3 (33)	100	33
R&D	3 (67)	0	3 (67)	0	33	33
Foreign Currency	2 (50)	0	0	2 (50)	50	0
EPS	1 (100)	0	1 (100)	0	100	0
Related Party Disclosures	1 (100)	0	0	1 (100)	100	0
Nonmonetary transactions	1 (100)	0	0	1 (100)	100	0
Other	5 (20)	1 (0)	2 (50)	2 (0)	60	80
Total	515 (44)	159 (38)	272 (52)	84 (29)	39	13

Areas are ordered according to the total frequency with which they appeared in Column 2.

[a] The 178 reserves include general reserves (102 attempts), restructuring reserves (29 attempts), reserves established in business combinations (37 attempts), and reserves against deferred-tax assets (10 attempts).

to those areas.[11] The instructions asked participants to give three experiences for the first area, but if they ran out of experiences for the first area they could move onto the second area, and so on, until they supplied a total of three experiences. This approach was designed to ensure that our sample included earnings management decisions pertinent to numerous accounting topics.

Twenty percent of surveys asked "no list" participants to provide a total of three experiences where companies attempted to materially manage earnings. "No list" responses were elicited to provide an estimate of the relative frequency of various attempts.

Survey Questions

Following the instructions (for all participants) and the table of five accounting areas (for list participants), the survey contained a sequence of 17 questions for "Attempted Earnings Management Experience #1," repeated those questions for Attempts #2 and #3, and finished with a short debriefing questionnaire. Included among the questions were requests that participants explain the attempt (called "the issue"), including the manner in which accounts were affected, and that participants state the specific criteria in financial accounting standards or principles that were most important for determining appropriate treatment of the attempt. Participants also stated their views of the manager's incentive(s) for the attempt.

Participants chose from among four alternatives that described how the attempt was eventually treated in the audited financial statements:

(1) "treated as the company originally desired, because they demonstrated they were complying with GAAP,"
(2) "treated as the company originally desired, because there was no convincing evidence that the company's position was incorrect,"
(3) "treated as the company originally desired, for some other reason (please explain)," and
(4) "treated other than how the company originally desired (please explain)."

Participants rated the materiality of the attempt to the client's financial reporting, with the alternatives of "Material by itself: (a) in current year (b) in future years," "Material in combination with other issues (a) in current year (b) in future years," and "Not material." They rated the client's annual net sales to provide an estimate of client size, with the alternatives of (a) < \$50M (million), (b) \$50M-\$100M, (c) \$100M–\$250M, (d) \$250M–\$1B (billion), and (e) > \$1B. They provided additional data, such as the year in which the attempt occurred, the industry of the client, the number of times they had encountered the attempt in their careers, and the percentage of times the attempt had been treated in the audited financial statements in a manner consistent with each of the alternatives listed above. The survey concluded with a short debriefing questionnaire. Participants returned the survey in an enclosed, stamped, addressed envelope.

Coding the Data

Descriptive Data

We provide descriptive data about areas in which earnings management was attempted, and auditors' assessments of managers' incentives for these attempts. To do so, one of the authors coded attempts according to financial accounting area (e.g., revenue recognition), incentive(s) for the attempt, and primary incentive for the attempt. When a participant listed multiple incentives and did not indicate one incentive as primary, the first incentive listed was coded as primary. To insure consistency of coding, all of the authors agreed on the coding scheme, the same author coded all attempts, a second author checked the first author's interpretation and coding, and the two met to resolve disagreements.

---

[11] Between participants we varied whether an area appeared on a table, and the area's 1–5 position, while insuring that each area appeared in tables an equal number of times and at each position within the table an equal number of times.

Variables Used in Hypothesis Tests

Participants supplied some data for hypothesis tests in the form of quantitative or categorical responses (e.g., client size, attempt materiality).[12] To test hypotheses about the effects of current-year income effect, precision of standard, and whether the transaction was structured, variables were coded independently by two coders as well as by one of the authors, based on auditors' descriptions of attempts, client incentives, and relevant financial accounting standards. The coders were CPAs, had public accounting experience of 5 and 1.5 years, respectively, and were unaware of hypotheses. Inter-coder agreement was high.[13] The results of all analyses are similar if based on any coder's individual codes. Following independent coding, coders and the coding author met to resolve disagreements and create a consensus coding for each variable. The results reported are based on consensus coding.

Precision of Standard. Some participants specifically identified the key criterion in financial accounting standards or principles that determined appropriate treatment of the attempt, (e.g., "FAS No. 13 capitalization criterion number 4 (90 percent PV lease payment))." When participants were less specific or did not respond to this question, each coder examined the description of the attempt and identified the key criterion affecting disposition of the attempt. If insufficient information existed to identify the key criterion, then the Precision score was coded as missing.

After the key criterion was identified, Precision was coded 1 if the criterion specifically allowed or disallowed the accounting treatment or if the criterion was quantified, and 0 otherwise (e.g., if the criterion was not specific, required estimation, or used vague verbal definitions rather than numerical definitions). For example, SFAS No. 5 specifically disallows accruing contingent gains or unspecified contingent losses, so it would be coded as precise with respect to attempts governed by that criterion. On the other hand, SFAS No. 5 states that accrual of a specific contingent loss depends on the loss being "probable" and "reasonably estimable," so it would be coded as imprecise with respect to attempts governed by that criterion. If multiple key criteria were present and at least one was imprecise, the Precision score was 0. Thus, APB No. 18 is coded as imprecise with respect to the "20 percent of ownership" threshold used to determine application of the equity method, because APB No. 18 also indicates that additional evidence can be used to determine whether the investor has "significant influence" over the investee. This procedure therefore biases toward coding standards as imprecise. In Section IV we discuss how this aspect of our coding procedure affects our inferences.

---

[12] Auditors identified whether the attempt was material individually and in combination with other issues. We developed two materiality measures: Matrate (= 1 if, in any year, the attempt was material individually or in combination with other issues, 0 if not), and Matcur (= 1 if, in the current year, the attempt was material individually or in combination with other issues, 0 if not). These two ratings are highly correlated (Pearson correlation = 0.93, $p < 0.0001$), and produce the same results when included in analyses, so results are only reported for Matrate. We also identified 83 attempts for which materiality was mentioned explicitly by auditors, and developed a Matsaid measure to capture mention (Matsaid = 1 if materiality mentioned, 0 if not). Auditors were more likely to mention materiality when an attempt was rated as immaterial (e.g., Pearson correlation between Matsaid and Matrate = –0.58, $p < 0.0001$).

[13] We used pairwise Kappa statistics to assess coder agreement with each other and with the author who also coded the data. For current-period-income effect, Kappas were 0.78 (coder1 to coder2), 0.80 (coder1 to author), and 0.77 (coder2 to author). For precision of standard, Kappas were 0.67 (coder1 to coder2), 0.74 (coder1 to author), and 0.68 (coder2 to author). For transaction structuring, Kappas were 0.71 (coder1 to coder2), 0.73 (coder1 to author), and 0.71 (coder2 to author). In general, agreement was high (Kappas of 0.8 indicate excellent agreement; Kappas of 0.4 indicate moderate agreement [Stokes et al. 2000]).

Transaction Structuring. Whenever an attempt involved a change in the timing or nature of a contract, transaction or activity, as opposed to involving a judgment or estimation process, Structuring was coded as 1, and 0 otherwise. Examples of structured transactions include "not shipping a product available for shipment because the quarter had already reached budgeted level" (modifying timing of a transaction), "writing contract for revolving credit facility to not require any minimum payments within 12 months to avoid current liability recognition, even though the majority of cash receipts are applied to the balance" (modifying contract), "employing a third-party guarantor of residual value to allow capital-lease treatment for lessor and operating-lease treatment for lessee" (modifying transaction), and "hastening delivery of materials to shop floor to recognize revenue under percentage of completion method" (modifying activity).

Current-Year Income Effect (CIE). CIE was coded as 1 (–1) if the attempt increased (decreased) current-year income. CIE was 0 if the attempt had no clear CIE (e.g., "accruing large reserves and offsetting them with goodwill in a purchase-method acquisition") or if the attempt as described did not provide sufficient information to identify the CIE (e.g., "Over-accrue reserve for receivable in profitable year and reverse in lean year" does not focus on a particular year).

IV. RESULTS

Descriptive Data about Attempts, Incentives, and Industries

Earnings Management Attempts

Column 2 of Table 2 lists attempts by the accounting areas included in the tables provided to list participants, ordered by the total frequency of responses.[14] On average, attempts occurred 2.7 years prior to data collection (median 2 years), and almost all occurred within five years of data collection. Similar to prior analyses of SEC enforcement releases, many attempts involved revenue recognition (78/515 = 15 percent), but the most common attempts in our sample involve contingencies and reserves (178/515 = 35 percent). Numerous attempts also involved business combinations, fixed assets, intangibles, investments, and leases.

Incentives

Descriptive data about auditors' beliefs about managers' incentives for earnings management attempts are shown in Table 3. Consistent with prior research (for a recent review, see Fields et al. [2001]), Panel A provides evidence that attempts were associated with stock market and contracting incentives. The incentive identified most often was meeting analyst expectations (86 attempts), and auditors required adjustment of these attempts 28 percent of the time. Overall, stock-market-related incentives account for 56 percent (150/266) of specific primary incentives.

Some participants identified a particular financial statement effect or pattern as the primary reason for an attempt. Consistent with prior research, Panel B of Table 3 provides evidence that attempts were intended to affect income statement numbers, to smooth income, and to improve future income. Thirty-three percent (78/238) refer to smoothing or managing a trend, while 25 percent (59/238) involve increasing future income.

---

[14] Comparisons (not shown) between list and no list respondents reveal very similar rankings of frequencies across the rows of Table 2. However, list respondents identified attempts in more areas, enhancing coverage of areas as intended. In all analyses, we combine list and no list responses.

TABLE 3
Auditors' Perceptions of Managers' Incentives for Attempting Earnings Management

	Number of Attempts	*Percent of Attempts Adjusted by Auditor*
Panel A: Incentives Identified as Primary Motivation for Attempt[a]		
Related to Stock Market		
Analyst expectations	86	28
Stock price	29	56
Impending or new IPO or SEO	12	55
Mention investors	8	75
Selling company	8	75
SEC registrant	5	40
Do because market doesn't detect	2	0
Total	150	41
Related to Contracting and Cash Flows		
Executive Compensation		
Bonuses/incentives	22	50
Board expectations	7	57
Total	29	52
Debt		
Debt covenants	16	69
Maintain access to credit	5	80
Leverage	4	25
Bond ratings	2	0
Total	27	59
Concerned about reaction of parent	17	41
Other cash effects (e.g., minimizing accounting cost, negotiating contracts)	11	45
Privacy (e.g., concerned about competition, confidentiality)	8	62
Regulatory requirements	7	57
Politics	4	75
Taxes	2	50
Total Contracting/Cash Flows	105	56
Clarity (e.g., trying to better reflect economics, avoid confusion)	11	36
Total Specific Incentives	266	47

(Continued on next page)

TABLE 3 (Continued)

	Number of Attempts	*Percent of Attempts Adjusted by Auditor*
Panel B: Particular Numbers/Patterns Identified as Primary Motivation for Attempt[b]		
Focus on Statements/Numbers		
Current income statement	47	51
Income target or projection	32	53
Avoiding loss	14	50
Current balance sheet	6	33
EBITDA	2	0
Total	101	49
Smoothing Income		
Smoothing	36	28
Trend/growth	16	50
Reserve for future	9	33
Spread over future periods	7	29
Conservative	4	25
Manage future earnings	3	33
Manage future expectations	2	50
Offset gains and losses	1	0
Total	78	33
Improve Future Income		
Increase future income	56	37
Load expenses into one year	3	33
Total	59	37
Total numbers/patterns	238	41
Other	11	43
Grand Total, Panels A and B	515	44

---

[a] Panel A lists the frequency with which various incentives were identified by auditors as a manager's primary motivation for an earnings management attempt.

[b] Panel B lists the frequency with which various financial statement numbers or patterns were identified by auditors as a manager's primary motivation for an earnings management attempt.

As noted by Fields et al. (2001), few studies examine whether combinations of incentives motivate attempts, focusing instead on proxies for one particular incentive. Our participants could mention more than one incentive for an attempt, so we can provide descriptive evidence on combinations of incentives. Results indicate that 24 percent of attempts for which a contracting incentive was mentioned also had a stock-market incentive mentioned. This was particularly apparent for attempts related to executive compensation, where 39 percent of attempts also had a stock-market incentive mentioned (often in the context of stock-based compensation plans).

Industries

Auditors identified the industry of the client for each attempt, and also indicated the proportion of time they had spent auditing clients in various industries. Absent industry effects, we would expect the proportion of attempts reported for each industry to equal the proportion of time auditors spent auditing clients in that industry. While this is generally the case, we find more attempts in the electronics industry than would be suggested by respondents' experience ($p < .001$), particularly attempts that involve bill-and-hold revenue-recognition problems and problems with inventory reserves, and fewer attempts by not-for-profits ($p < .001$).

Hypothesis Tests

Paralleling Section II, we first discuss analyses of precision and structuring, then consider current-year income effects (CIE) and the effects of client size and materiality. Table 1 lists hypotheses and results.

Precision of Standards and Transaction Structuring

Table 4, Panel A shows a 2 × 2 contingency table of managers' earnings management attempts, crossing precision and structuring. Panel B shows the results of a loglinear analysis of the effect of precision and structuring on the relative frequency of attempts.[15] Consistent with H1, there is a significant interaction between precision and structuring, with 69 percent (46/67) of structured transactions associated with precise standards (Chi-square = 8.86; $p < .0029$), and 35 percent (156/448) of unstructured transactions associated with precise standards (Chi-square = 39.96; $p < .0001$).

Table 5, Panel A shows the percentage of managers' attempts that auditors adjusted, by precision and structuring.[16] Panel B shows the results of a logistic regression of adjust (EMA adjusted = 1; attempt not adjusted = 0) on precision, structuring, the precision × structuring interaction, materiality and size (no other interactions are significant).[17] Consistent with H2, the precision × structuring interaction is significant ($p < .0024$).[18]

Comparing cells of Table 5, Panel A clarifies these results. Structured attempts to manage earnings, whether against precise or imprecise standards, were usually successful (i.e., 21 percent were adjusted).[19] In fact, in ten instances auditors noted that they had helped structure the transaction to ensure compliance with GAAP. Unstructured attempts

---

[15] Analyses of managers' attempt decisions were performed via loglinear models using Proc Catmod (SAS v. 8). As described by Stokes et al. (2000, chapter 16), loglinear analyses treat all variables as response variables and test for statistical dependence and independence (analogous to correlation analysis). Loglinear models make use of the fact that statistical independence can be expressed in terms of a linear combination of the logarithms of the cell probabilities in a contingency table. A low p-value associated with a variable indicates a high probability that cell probabilities are not independent of that variable.

[16] Of 515 total attempts, 226 were adjusted. 110 attempts were not adjusted because the auditor believed the client had demonstrated compliance with GAAP, 88 because the auditor did not have convincing evidence that the client's position was incorrect, and 91 for some other reason. For 66 of those 91, the stated reason was immateriality. We use a binary adjustment variable in analyses because a finer distinction among adjustment decisions is subjective and therefore difficult to interpret. For example, an auditor faced with a potentially overstated bad debt expense might justify no adjustment by saying that the reserve was permitted under GAAP, that the auditor accepted management's estimate, or that the income effect was immaterial.

[17] Analyses of auditors' adjustment decisions were performed using Proc Logistic (SAS v. 8).

[18] Two simple effects in the precision × structuring interaction are significant. First, when standards are precise, the proportion of attempts adjusted when transactions are not structured (62 percent) exceeds the proportion adjusted when transactions are structured (15 percent; Chi-square = 25.06; $p < 0.0001$). Second, when transactions are not structured, the proportion of attempts adjusted when standards are precise (62 percent) exceeds the proportion adjusted when standards are imprecise (39 percent; Chi-square = 20.72; $p < 0.0001$).

[19] Imprecise standards (by our coding criteria) may exhibit a relatively high level of precision. For example, "bill and hold" revenue recognition criteria from AAER No. 108 and cost/equity method criteria from APB No. 18 have an element of imprecision, so were coded as precision = 0, but still provide opportunities for structuring transactions.

TABLE 4
Attempts to Manage Earnings

Panel A: Manager Attempts by Precision of Accounting Standard and Transaction Structuring

	Transaction Not Structured	*Transaction Structured*	*Total*
Low Precision	292	21	313
High Precision	156	46	202
Total	448	67	515

Panel B: Manager Attempts: Loglinear Analysis of the Effects of Precision and Structuring[a]

Maximum Likelihood Analysis of Variance

Source	*Chi-Square*	*Pr > ChiSq*
precision	0.31	0.5764
structuring	187.51	< 0.0001
precision × structuring	25.14	< 0.0001**

** $p < 0.05$ with respect to hypothesis test.

[a] The loglinear analysis in Panel B examines how the frequencies with which attempts appear in the cells of Panel A are influenced by precision and structuring (for further discussion of loglinear models, see Stokes et al. [2000]).
Precision = 1 if standard precise, 0 if standard not precise; and
Structuring = 1 if transaction structured, 0 if transaction not structured.

governed by imprecise standards also were usually successful (with 39 percent adjusted). However, unstructured attempts governed by precise standards were usually unsuccessful (with 62 percent adjusted). Examples include "treating as sales normal inventory that had not been shipped before year end," "setting up unsupported accruals for unspecified future losses," and "issuing stock options to employees below market price and not recording compensation expense." Auditors' comments indicated that these attempts often reflected lack of knowledge on the part of managers about the relevant accounting standards.

Current-Year Income Effect

Before adding current-year income effect (CIE) to the analysis, we drop observations that lack a clear CIE (i.e., for which CIE = 0), reducing our sample to 431 attempts. Table 6, Panel A shows managers' attempts in a 2 × 2 × 2 (precision × structuring × CIE) classification table. Panel B shows the results of a loglinear analysis of the effect of precision, structuring, and CIE on the relative frequency of attempts. Consistent with H1, precision interacts with structuring. Consistent with H3, we find that 63 percent (272/431) of attempts increase income.[20] Consistent with H4, managers are more likely to decrease income with attempts governed by imprecise standards. Consistent with H5, managers are more likely to decrease income with attempts that are not structured. Overall, these results

---

[20] The significance of CIE in the loglinear analysis is incremental to the significance of the CIE × structuring and CIE × precision interactions. However, a test of equal proportions indicates that the proportion of CIEs that are income increasing (63 percent) differs significantly from 50 percent (Chi-square = 29.63; $p < 0.0001$).

TABLE 5
Auditors' Adjustment Decisions

Panel A: Percentage of Attempts That Were Adjusted by the Auditor, by Precision of Accounting Standard and Transaction Structuring[a]

	Transaction Not Structured	*Transaction Structured*	*Total*
Low Precision	115/292 = 39	7/21 = 33	122/313 = 39
High Precision	97/156 = 62	7/46 = 15	104/202 = 51
Total	212/448 = 47	14/67 = 21	226/515 = 44

Panel B: Auditor Adjustment Decision: Logit of Precision, Structuring, Materiality, and Client Size[b]

Parameter	*Estimate*	*Standard Error*	*Wald Chi-Square*	*Pr > ChiSq*
Intercept	−1.4123	0.3731	14.3312	0.0002
precision	1.2306	0.2389	26.5366	< 0.0001
structuring	−0.4669	0.4932	0.8961	0.3438
precision × structuring	−2.0674	0.6800	9.2431	0.0024**
materiality	1.9698	0.2835	48.2882	< 0.0001**
size	−0.1957	0.0707	7.6554	0.0057**

** $p < 0.05$ with respect to hypothesis test.

[a] Cell formats are (number of attempts adjusted/total number of attempts = %).

[b] The dependent variable in the logit is adjust (auditor required adjustment of the attempt = 1, auditor did not require adjustment of the attempt = 0).

Precision = 1 if standard precise, 0 if standard not precise;
Structuring = 1 if transaction structured, 0 if transaction not structured;
Materiality = 1 if material in current or future year individually or in combination with another issue, 0 if never material; and
Size = 1 if annual sales < \$50M, 2 if \$50M < sales < \$100M, 3 if \$100M < sales < \$250M, 4 if \$250M < sales < \$1B, 5 if sales > \$1B.

suggest that managers are more likely in general to make earnings management attempts that increase current income, but are more likely to decrease current income with unstructured transactions and with transactions that involve imprecise standards.

Table 7, Panel A shows the percentage of attempts adjusted by auditors, by CIE, precision, and structuring. Panel B contains a reduced logit model that omits insignificant interactions.[21] Consistent with H2, precision interacts with structuring. Consistent with H6, auditors are significantly more likely to adjust attempts that increase current-year income.

---

[21] Pearson and Spearman correlations indicate some collinearity ($r < 0.25$ but $p < 0.05$) between independent variables (particularly between precision, structure and the precision × structuring interaction, and between size, CIE and materiality). Collinearity often decreases the significance of variables and therefore biases away from supporting hypotheses, but it could create unstable parameter estimates. However, all parameters have similar signs and magnitudes regardless of the other variables included in models, indicating that parameters are stable (Neter et al. 1983).

TABLE 6
Attempts to Manage Earnings by Current-Period Income Effect

Panel A: Manager Attempts to Manage Earnings, by Current-Year Income Effect, Precision of Accounting Standard, and Transaction Structuring

	Transaction Not Structured		*Transaction Structured*		
	Low Precision	*High Precision*	*Low Precision*	*High Precision*	*Total*
Income-Increasing	136	91	16	29	272
Income-Decreasing	113	42	0	4	159
Total	249	133	16	33	431

Panel B: Manager Attempts: Loglinear Analysis of the Effects of Current-Year Income Effect, Precision and Structuring[a]

Maximum Likelihood Analysis of Variance

Source	*Chi-Square*	*Pr > ChiSq*
CIE	15.61	< 0.0001**
precision	1.20	0.2726
structuring	58.71	< 0.0001
precision × structuring	8.61	0.0033**
CIE × precision	6.78	0.0092**
CIE × structuring	4.57	0.0325**

** $p < 0.05$ with respect to hypothesis test.

[a] The loglinear analysis in Panel B examines how the frequencies with which attempts appear in the cells of Panel A are influenced by precision, structuring and CIE (for further discussion of loglinear models, see Stokes et al. [2000]).

CIE = 1 if current-year-income increasing, –1 if current-year-income decreasing;
Precision = 1 if standard precise, 0 if standard not precise; and
Structuring = 1 if transaction structured, 0 if transaction not structured.

Materiality and Client Size

Analyses in Tables 5 and 7 indicate that, consistent with H7, auditors are more likely to adjust attempts they label as material. When auditors considered an attempt to be material, it was adjusted 53 percent (199/376) of the time; when auditors considered an attempt to be immaterial, it was adjusted 17 percent (22/130) of the time (materiality ratings were not available for nine observations). These results must be interpreted with caution, because we do not know the dollar amounts of attempts or auditors' quantitative materiality thresholds. However, given that participants were asked to describe "experiences where companies attempted to materially manage earnings," they probably believed at some point that the attempts were material. To the extent that attempts that are not adjusted due to immateriality actually are material, this result is consistent with regulatory concern (e.g., Levitt

TABLE 7
Auditors' Adjustment Decisions by Current-Period Income Effect

Panel A: Percentage of Attempts That Were Adjusted by the Auditor, by Current-Year Income Effect, Precision of Accounting Standard, and Transaction Structuring[a]

	Transaction Not Structured		*Transaction Structured*		
	Low Precision	*High Precision*	*Low Precision*	*High Precision*	*Total*
Income-Increasing	63/136 = 46	65/91 = 71	7/16 = 44	7/29 = 24	142/272 = 52
Income-Decreasing	43/113 = 38	17/42 = 40	0/0	0/4 = 0	60/159 = 38
Total	106/249 = 42	82/133 = 61	7/16 = 44	7/33 = 21	202/431 = 47

Panel B: Auditor Adjustment Decision: Logit of Precision, Structuring, CIE, Materiality, and Client Size[b]

Parameter	*Estimate*	*Standard Error*	*Wald Chi-Square*	*Pr > ChiSq*
Intercept	−1.1910	0.3964	9.0283	0.0027
precision	1.0025	0.2573	15.1793	< .0001
structuring	−0.2100	0.5513	0.1451	0.7033
precision × structuring	−1.8772	0.7358	6.5080	0.0107**
CIE	0.2595	0.1162	4.9895	0.0255**
materiality	1.7239	0.2965	33.8092	< .0001**
size	−0.1811	0.0759	5.6894	0.0171**

** $p < 0.05$ with respect to hypothesis test.

[a] Cell formats are (number of attempts adjusted/total number of attempts = %).

[b] The dependent variable in the logit is adjust (auditor required adjustment of the attempt = 1, auditor did not require adjustment of the attempt = 0).

CIE = 1 if current-year-income increasing; −1 if current-year-income decreasing;
Precision = 1 if standard precise, 0 if standard not precise;
Structuring = 1 if transaction structured, 0 if transaction not structured;
Materiality = 1 if material in current or future year individually or in combination with another issue, 0 if never material; and
Size = 1 if annual sales < \$50M, 2 if \$50M < sales < \$100M, 3 if \$100M < sales < \$250M, 4 if \$250M < sales < \$1B, 5 if sales > \$1B.

1998) and recent experimental findings (Braun 2001; Libby and Kinney 2000) that auditors sometimes justify not adjusting potentially material attempts by labeling them as immaterial.

Analyses in Tables 5 and 7 also indicate that, consistent with H8, auditors are more likely to adjust their smallest clients' attempts. Attempts made by the smallest clients (annual net sales of \$50 million or less) were adjusted 62 percent (56/91) of the time, while attempts made by the largest clients (annual net sales of \$1 billion or greater) were adjusted

28 percent (38/134) of the time.[22] This result may indicate that larger clients have more accounting sophistication, or that auditors are more willing to allow earnings management by their most important clients.

Adjustment Decisions, by Financial Accounting Area

The results of our analyses indicate that precision of standards, transaction structuring, and CIE are related to attempt and adjustment decisions. Table 2 presents the percentage of attempts adjusted, by accounting area and CIE, and also lists the percent of attempts in each area that were associated with a precise standard and that were structured. Several patterns in Table 2 are consistent with the results of our hypothesis tests. First, attempts involving reserves are likely to decrease current income, and attempts involving reserves and intangibles are likely to be unstructured and to involve imprecise standards, consistent with their relatively low adjustment percentage (35 percent and 44 percent, respectively). Attempts involving leases, compensation, and consolidations and equity vs. cost method are likely to be structured against precise standards, consistent with their low adjustment percentage (39 percent, 16 percent, and 27 percent, respectively). Attempts involving revenue-recognition, investments, and transfers of receivables tend to increase current-period income. Therefore, while these attempts are made frequently, they are more often than not adjusted. Finally, further analyses (not shown) indicate that non-reserve business combination attempts are usually not adjusted when they decrease current income (e.g., expensing in-process R&D) or when they are structured around precise standards (e.g., to enable pooling), but are usually adjusted otherwise, consistent with the effect of CIE and the precision × structuring interaction identified in analyses of auditors' adjustment decisions. Fixed-asset attempts are difficult to reconcile with our analyses, since they are more likely to be adjusted when they decrease current-year income.

V. SUMMARY, IMPLICATIONS, LIMITATIONS, AND DIRECTIONS FOR FUTURE RESEARCH

In this study, we analyze auditors' recalled experiences with clients who they believe are attempting to manage earnings. This approach allows us to identify earnings management attempts for description and analysis, to separate managers' attempt decisions from auditors' decisions about whether to require adjustment of attempts, and to examine the effects of general factors on these decisions. Our sample contains earnings management attempts in a variety of accounting areas (most frequently, with respect to reserves), motivated by a variety of incentives (most frequently, to affect the stock market), and in a variety of industries.

Our results indicate that precision and structuring interact in managers' and auditors' decisions: when transaction structuring is involved, managers are more likely to make attempts (and auditors are less likely to adjust attempts) that are governed by precise standards, and when transaction structuring is not involved, managers are more likely to make attempts (and auditors are less likely to adjust attempts) that are governed by imprecise standards. Managers tend to make attempts that increase current-year income, but auditors are more likely to adjust these attempts. Managers are more likely to make attempts that decrease current-year income when standards are imprecise and/or with unstructured transactions. Auditors are more likely to adjust attempts they consider material and attempts made by small clients.

---

[22] Additional analyses (not shown) indicate a size effect when the sample is restricted to include only material, income-increasing, non-GAAP attempts. Specifically, when attempted by clients of the smallest size category (annual net sales of $50 million or less), 88 percent (37/42) were adjusted; when attempted by clients of the largest size category (annual net sales of $1 billion or greater), 54 percent (15/28) were adjusted.

Our results shed light on the auditor's role in reducing earnings management. Most prior studies infer earnings management from issued financial reports, and therefore cannot address the effectiveness of the audit function in thwarting earnings management attempts. Many attempts in our sample were adjusted by auditors, particularly when the attempt would have increased current-year income. Our results probably understate auditors' role, as we have no way of assessing the extent to which knowledge of an impending audit deters managers from even attempting earnings management.

Second, our results highlight general factors that affect where earnings management is attempted and left unadjusted. This evidence may be of interest to regulators and standard setters. Our results imply that managers respond to increased precision by structuring attempts more often and by making fewer income-decreasing attempts. Our results also imply that auditors respond to increased precision by adjusting unstructured attempts more often and structured attempts less often. Consistent with regulatory concerns, our results indicate that auditors are relatively less likely to adjust current-period-income-decreasing attempts, even though such attempts may allow managers to increase income in a future period, and that auditors are less likely to adjust attempts made by their larger clients.

Our study is subject to limitations. First, the validity of our results depends on the accuracy of our survey data. Some earnings management attempts may be underrepresented because they are experienced more by auditors who did not respond to our survey, or because they are difficult for auditors to detect or recall. Other attempts could be overrepresented because they are salient (e.g., those mentioned in press coverage of former SEC Chairman Levitt's "Numbers Game" speech a month prior to data collection). A related concern is self-serving responses, in which auditors recall having behaved more appropriately than they actually did. Any such self-serving bias would tend to understate the proportion of attempts reported as not adjusted, particularly those not adjusted for reasons other than compliance with GAAP. Therefore, our evidence that 56 percent of attempts were not adjusted may understate the extent to which detected attempts are not adjusted in practice. In general, these concerns are less likely to affect the analyses of general factors (particularly interactions) than they are to affect the frequency of recalled individual attempts or adjustment decisions.

A second limitation is that audit-client-confidentiality conditions, which prohibit us from knowing the identities of the sample companies, preclude linking our results to stock market performance or to firm-specific or market variables that we did not elicit. In addition, our sample includes only earnings management attempts, so we cannot examine factors affecting the likelihood managers will attempt earnings management or auditors will detect an attempt—our results are conditioned on earnings management having been identified and recalled by the auditor. Finally, we cannot assess the extent to which the presence of auditors deters managers from attempting earnings management.

Our results suggest opportunities for future research. Because our data provide a baseline of managers' and auditors' behavior prior to the recent economic downturn and flurry of earnings-management-related speeches, SABs, and SASs, a future replication could refer to this baseline to assess change. We would expect that the relative proportions of various types of attempts might vary in different contexts, but the directional significance of main effects and interactions would be robust to changes in context. Our results could guide creation of auditor training materials or decision aids that help auditors identify attempts and consider whether they should be adjusted. The relative frequencies of attempt areas and incentives that we document may better focus archival and experimental researchers on the areas in which earnings management attempts are most frequent.

APPENDIX

INSTRUCTIONS

[Note: "No list" participants received a version that included only the unbracketed text. "List" participants received a version that replaced the "In the pages that follow..." sentence with the bracketed text, and also received a table listing five accounting issues and related financial accounting standards.]

The purpose of this study is to learn about your experiences with companies attempting material "earnings management." We interpret "earnings management" broadly to include all efforts to influence accounting measurement, recognition, and/or disclosure that are intended to affect the view of an entity obtained by a person using the financial statements. We interpret "material" to indicate an amount that could affect financial statement users' judgments. We focus on *attempted* earnings management to include all experiences where the company attempted to manage earnings, whether or not the final audited financial statements appeared as the company initially desired.

Since audited financial statements must conform with GAAP, companies who wish to manage earnings must do so within the constraints of the relevant financial accounting standards. Ways in which earnings management can be attempted include:

- Structuring transactions in a manner that gets around a standard to achieve a desired accounting effect.
- Aggressively interpreting the wording of a standard in a way that justifies a desired accounting effect.
- Aggressively interpreting the data associated with transactions in such a way that, when compared to the standard, it justifies a desired accounting effect.

In the pages that follow, we ask you to list a total of three experiences where companies attempted to materially manage earnings.

[In Table 1 on the next page, we list five broad accounting topics that have been addressed by accounting standards. In the pages that follow Table 1, we ask you to list a total of three experiences where companies attempted to materially manage earnings with respect to those topics. The first topic is the one for which we most need to learn about your experiences, so if possible, we would like you to list three experiences with attempted earnings management that are related to Topic 1. However, it may be that you have fewer than three experiences with attempted earnings management that are related to Topic 1. If so, please document those Topic 1 experiences (if any) that you have had, and turn to Topic 2. If documenting your Topic 2 experiences allows you to reach a total of three experiences across Topics 1 and 2, stop; if not, please turn to Topic 3, and so on.]

In all cases, please focus on those experiences that you feel are most important, in the sense of being the most frequently occurring examples of attempted material earnings management that you have encountered. Please consider all of your experiences of this nature, regardless of (1) whether or not the transaction was eventually treated in the manner preferred by the company, and (2) whether standards have been revised subsequently to more effectively address that sort of earnings management.

Company and auditor confidentiality is assured because we are not asking for the identity of companies or auditors. Your responses will be anonymous. **This study has been cleared with [firm name] to be sure that Firm policies and professional standards are upheld, and to be sure that the results of this project are very useful to [firm name].**

REFERENCES

Antle, R., and B. Nalebuff. 1991. Conservatism and auditor-client negotiations. *Journal of Accounting Research* 29 (Supplement): 31–59.

Arens, A., and J. Loebbecke. 1997. *Auditing: An Integrated Approach*. 7th edition. Englewood Cliffs, NJ: Prentice Hall.

Ayers, S., and S. E. Kaplan. 1998. Potential differences between engagement and risk review partners and their effect on client acceptance judgments. *Accounting Horizons* 12 (2): 139–153.

Barth, M. E., J. A. Elliott, and M. W. Finn. 1999. Market rewards associated with patterns of increasing earnings. *Journal of Accounting Research* 37 (Autumn): 387–413.

Beatty, A., B. Ke, and K. Petroni. 2002. Earnings management to avoid earnings declines across publicly and privately held banks. *The Accounting Review* (July): 515–546.

Beneish, M. D. 1997. Detecting GAAP violation: Implications for assessing earnings management among firms with extreme financial performance. *Journal of Accounting and Public Policy* 16 (Fall): 271–309.

Beresford, D. R. 1999. It's time to simplify accounting standards. *Journal of Accountancy* 187: 65–67.

Bonner, S. E., Z-V. Palmrose, and S. M. Young. 1998. Fraud type and auditor litigation: An analysis of SEC accounting and auditing enforcement releases. *The Accounting Review* 73 (October): 503–532.

Bradshaw, M., S. A. Richardson, and R. G. Sloan. 2001. Do analysts and auditors use information in accruals? *Journal of Accounting Research* 39 (1): 45–74.

Braun, K. W. 2001. The disposition of audit-detected misstatements: An examination of risk and reward factors and aggregation effects. *Contemporary Accounting Research* 18 (1): 71–100.

Burgstahler, D., and I. Dichev. 1997. Earnings management to avoid earnings decreases and losses. *Journal of Accounting and Economics* 24: 99–126.

———, and M. Eames. 1998. Management of earnings and analyst forecasts. Working paper, University of Washington and Santa Clara University.

Carcello, J. V., and Z-V. Palmrose. 1994. Auditor litigation and modified reporting on bankrupt clients. *Journal of Accounting Research* 32 (Supplement): 1–29.

Davies, G. M. 1993. Witnessing events. In *Memory in Everyday Life: Advances in Psychology*, edited by G. M. Davies, and R. H. Logie, 367–401. Amsterdam, The Netherlands: North-Holland/Elsevier Science Publishers.

Dechow, P. M., R. G. Sloan, and A. P. Sweeney. 1995. Detecting earnings management. *The Accounting Review* 70 (April): 193–225.

———, ———, and ———. 1996. Causes and consequences of earnings manipulation: An analysis of firms subject to enforcement actions by the SEC. *Contemporary Accounting Research* 13: 1–36.

———, and D. J. Skinner. 2000. Earnings management: Reconciling the views of accounting academics, practitioners, and regulators. *Accounting Horizons* 14 (June): 235–250.

DeFond, M. L., and J. Jiambalvo. 1993. Factors related to methods. *Contemporary Accounting Research* 9: 415–431.

Degeorge, F., J. Patel, and R. Zeckhauser. 1999. Earnings management to exceed thresholds. *Journal of Business* 72: 1–33.

Diver, C. S. 1983. The optimal precision of administrative rules. *Yale Law Journal* 93: 65–109.

Ehrlich, I., and R. A. Posner. 1974. An economic analysis of legal rulemaking. *Journal of Legal Studies* 3: 257–286.

Fields, T., T. Lys, and L. Vincent. 2001. Empirical research on accounting choice. *Journal of Accounting and Economics* 31 (1-3): 255–307.

Flanagan, J. C. 1954. The critical incident technique. *Psychological Bulletin* 51 (4): 327–358.

Francis, J., and D. Simon. 1987. A test of audit planning in the small-client segment of the U.S. Audit Market. *The Accounting Review* 62 (January): 145–157.

———, and J. Krishnan. 1999. Accounting accruals and auditor reporting conservatism. *Contemporary Accounting Research* 16 (Spring): 135–165.

Gibbins, M., and J. Newton. 1994. An empirical exploration of complex accountability in public accounting. *Journal of Accounting Research* 32(2) (Autumn): 165–186.

———, S. Salterio, and A. Webb. 2001. Evidence about auditor-client management negotiation concerning client's financial reporting. *Journal of Accounting Research* 39 (3): 534–563.

———. 2001. Incorporating context into the study of judgment and expertise in public accounting. *International Journal of Auditing* 5 (3): 225–236.

Gibson, A. M., and A. H. Frakes. 1997. Truth or consequences: A study of critical issues and decision making in accounting. *Journal of Business Ethics* 16 (2): 161–171.

Gifford, D. J. 1971. Communication of legal standards, policy development and effective conduct regulation. *Cornell Law Review* 56: 409–468.

Graham, J. R., and C. R. Harvey. 2001. The theory and practice of corporate finance: Evidence from the field. *Journal of Financial Economics* 61 (2,3): 187–243.

Greenberg, B. G., R. R. Kuebler, Jr., J. R. Abernathy, and D. G. Horvitz. 1971. Application of the randomized response technique in obtaining quantitative data. *Journal of the American Statistical Association* 66: 243–250.

Hackenbrack, K., and M. W. Nelson. 1996. Auditors' incentives and their application of financial accounting standards. *The Accounting Review* 71 (January): 43–59.

Hayn, C. 1995. The information content of losses. *Journal of Accounting and Economics* 20: 125–153.

Healy, P. M., and J. M. Wahlen. 1999. A review of the earnings management literature and its implications for standard setting. *Accounting Horizons* 13 (October): 365–383.

Icerman, R., and W. Hillison. 1991. Disposition of audit-detected errors: Some evidence on evaluative materiality. *Auditing: A Journal of Practice & Theory* 10 (Spring): 22–34.

Jones, J. J. 1991. Earnings management during import relief investigations. *Journal of Accounting Research* 29 (Autumn): 193–228.

Kadous, K., J. Kennedy, and M. E. Peecher. 2000. Auditors' judgments of the acceptability of client-preferred accounting methods. Working paper, University of Washington, WA.

Kaplow, L. 1999. General characteristics of rules. In *Encyclopedia of Law and Economics*, edited by B. Bouckaert, and G. De Geest, 502–528. Cheltenham, U.K.: Edward Elgar.

Kennedy, J., D. N. Kleinmuntz, and M. E. Peecher. 1997. Determinants of the justifiability of performance in ill-structured audit tasks. *Journal of Accounting Research* 35 (Supplement): 105–123.

Kessler, D., and M. McClellan. 1996. Do doctors practice defensive medicine? *Quarterly Journal of Economics* 111: 352–390.

King, R. R. 2002. An experimental investigation of self-serving biases in an auditing trust game: The effect of group affiliation. *The Accounting Review* 77 (2): 265–284.

Kinney, W. R., and R. D. Martin. 1994. Does auditing reduce bias in financial reporting? A review of audit-related adjustment studies. *Auditing: A Journal of Practice & Theory* 13 (1): 149–156.

Lee, C-W. J., and Z. Gu. 1998. Low balling, legal liability and auditor independence. *The Accounting Review* 73 (October): 533–556.

Levitt, A. 1998. The numbers game. Speech delivered at the NYU Center for Law and Business, New York, NY, September 28.

Libby, R., and W. R. Kinney. 2000. Earnings management, audit differences, and analysts' forecasts. *The Accounting Review* 75 (4): 383–404).

———, R. Bloomfield, and M. W. Nelson. 2001. Experimental research in financial accounting. *Accounting, Organizations and Society* 27 (8): 775–810.

Loftus, E. F. 1982. Memory and its distortions. In *The G. Stanley Hall Lecture Series, Vol 2*, edited by Kraut, 123–154. Washington, D. C.: American Psychological Association.

Mason, A. K., and M. Gibbins. 1991. Judgment and U.S. accounting standards. *Accounting Horizons* 5: 14–24.

Mayhew, B. W., J. W. Schatzberg, and G. R. Sevcik. 2001. The effect of accounting uncertainty and auditor reputation on auditor objectivity. *Auditing: A Journal of Practice & Theory* 20: 49–70.

McNichols, M., and G. P. Wilson. 1988. Evidence of earnings management from the provision for bad debts. *Journal of Accounting Research* 26 (Supplement): 1–31.

———. 2000. Research design issues in earnings management studies. *Journal of Accounting and Public Policy* 19 (4/5): 313–345.

Miller, M. W., and L. Berton. 1993. Softer numbers: As IBM's woes grew, its accounting tactics got less conservative. *Wall Street Journal* (April 7).

Moore, J. S., and A. K. Reichert. 1983. An analysis of the financial management techniques currently employed by large U. S. corporations. *Journal of Business Finance and Accounting* 10: 623–645.

Nelson, M. W., and W. R. Kinney. 1997. The effect of ambiguity on auditors' loss contingency reporting judgments. *The Accounting Review* 72 (April): 257–274.

Neter, J., W. Wasserman, and M. H. Kutner. 1983. *Applied Linear Regression Models*. Homewood, IL: Irwin.

Ogus, A. I. 1992. Information, error costs and regulation. *International Review of Law and Economics* 12: 411–421.

Oppenheim, A. N. 1992. *Questionnaire Design, Interviewing, and Attitude Measurement*. New York, NY: Pinter Publishers.

Panel on Audit Effectiveness. 2000. *Report and Recommendations Exposure Draft* (May 31). Stamford, CT: Public Oversight Board.

Parfet, W. U. 2000. Accounting subjectivity and earnings management: A preparer perspective. *Accounting Horizons 14*: 481–488.

Pulliam, S. 1988. Besting FAS-13. *Corporate Finance* (December): 31.

Randall, D. M., and A. M. Gibson. 1990. Methodology in business ethics research: A review and critical assessment. *Journal of Business Ethics* 9: 457–471.

Salterio, S., and L. Koonce. 1997. The persuasiveness of audit evidence: The case of accounting policy decisions. *Accounting, Organizations and Society* 22 (6): 573–587.

Schipper, K. 1989. Commentary on earnings management. *Accounting Horizons* 3 (December): 91–102.

Shavell, S. 1987. *Economic Analysis of Accident Law*. Cambridge, MA: Harvard University Press.

St. Pierre, K., and J. A. Anderson. 1984. An analysis of the factors associated with lawsuits against public accountants. *The Accounting Review* 59 (April): 242–263.

Stokes, M. E., C. S. Davis, and G. G. Koch. 2000. *Categorical Data Analysis Using the SAS System*. Cary, NC: SAS Institute, Inc.

Trahan, E. A., and L. J. Gitman. 1995. Bridging the theory-practice gap in corporate finance: A survey of chief financial officers. *The Quarterly Review of Economics and Finance* 35 (1): 73–87.

Trompeter, G. 1994. The effect of partner compensation schemes and generally accepted accounting principles on audit partner judgment. *Auditing: A Journal of Practice & Theory* 13 (2): 56–71.

Wallace, R., and C. Mellor. 1988. Nonresponse bias in mail accounting surveys: A pedagogical note. *British Accounting Review* 20: 131–139.

Warner, S. L. 1965. Randomized response: A survey technique for eliminating evasive answer bias. *Journal of the American Statistical Association* 60: 63–69.

Wright, A., and S. Wright. 1997. An examination of factors affecting the decision to waive audit adjustments. *Journal of Accounting, Auditing & Finance* 12 (1): 15–36.

Zhang, P. 1999. A bargaining model of auditor reporting. *Contemporary Accounting Research* 16 (Spring): 167–184.

THE ACCOUNTING REVIEW
Vol. 77
Supplement 2002
pp. 203–211

DISCUSSION OF Evidence from Auditors about Managers' and Auditors' Earnings Management Decisions

Michael Gibbins
University of Alberta

Nelson et al. (2002) (hereafter the paper) consider how auditors respond to clients' attempts to manage earnings. Their research approach, using a field questionnaire to ask auditors to describe cases of attempted earnings management and to do some self-analysis of their cases, is both laudable and problematic. It is laudable because it can reveal important features of the auditor-client relationship, and it is problematic because the auditors' self-reporting may be subject to error and bias and may require significant interpretation. As an author of such research (e.g., Gibbins and Newton 1994; Gibbins et al. 2001; Gibbins and Trotman 2002), I support such work and support the paper. My discussion, therefore, is intended to provide a context for understanding the paper's contribution. The discussion begins with how the authors view earnings management, then turns to method and results.

I. THE PAPER'S THEORY ABOUT EARNINGS MANAGEMENT

The paper focuses on earnings management attempts and auditor responses to those attempts. The authors develop their hypotheses using four theory planks:

1. Management's intent in attempting to manage earnings (effect on current year's income);
2. Implemented either by structuring transactions to suit the accounting standards (a costly mechanism) or by convincing the auditor that management's desired accounting treatment reflects a correct interpretation of the rules (less costly);
3. Ease of implementation of the mechanism depending on the precision of relevant accounting standards;
4. Success in achieving earnings management depending on whether the auditor opposes management's attempt by requiring an adjustment to thwart it.

Management Intent

A major claim in the paper is that the questionnaire elicitation of *attempts* to manage earnings moves the examination of earnings management from the financial statement results back toward the client managers' intent. The auditor's role is to remove or reduce such attempts by requiring adjustments, so actual instances of earnings management are only a subset of the total attempts to manage earnings, the subset not thwarted by the auditor. Attempts are thus a clearer indication of management intent than is successful earnings management.

Management's intent is not benign: both attempts and successes at earnings management are undesirable, at least for people other than the managers. Earnings management hurts quality of earnings; it never helps. The auditor's job is to oppose, not to assist. This view is consistent with current worries about financial accounting manipulation, and with the definition in Schipper (1989), referred to at the beginning of the paper, that managers intend some private gain.

Current-Year Income Effect

The paper represents observable intent by "current-year income effect." The paper proposes that increasing current income is the primary management intent. It is a function of a proposed management wish to have "smoothly increasing income" over time. This central part of the theory is only briefly supported. It, and H3 that follows from it, are part of the paper's definition of earnings management, rather than being theoretically argued.

The paper states that since smoothly increasing income is the basic intent, management would want to decrease current income only in order to enable future income-increasing attempts. There is no consideration of the possibility that some managers may be conservative and choose income-reducing accounting policies, nor is a smoothing strategy distinguished by the paper's argument. A smoothing strategy could require increasing or decreasing current income in a given year: if income is to be smoothly increasing, then it would be pushed up to or down to the increasing trend line each year. The paper does not distinguish pushing up from a general wish to increase income, but does argue that pushing down must mean a wish to increase income in the future. The trade-off between current and future income is also a function of accrual accounting; some earnings management, such as recording over-generous asset impairments or writing off assets in a year of poor performance, is at least partly done to improve future earnings automatically by moving future expenses into the present.

Looking to the future, the paper suggests that the flexibility to undertake future earnings management would be facilitated by imprecise rules and by not structuring transactions. There would be room to maneuver. Like the assumption that earnings management attempts will favor current earnings, the statement about future flexibility is presented as largely self-evident. It is reasonable, but the position that no matter what direction an earnings management attempt takes, it must be to increase income, now or later, is an assertion rather than a theory.

Precision in Standards

Perhaps the most interesting feature of the paper is its proposal that precision in accounting standards affects how earnings are managed. My reading of the idea is that a precise standard provides little scope for opportunistic interpretations, leading the managers to choose structuring transactions as the way to accomplish their earnings management goals. In addition, a precise standard aids this type of earnings management, because managers can structure directly to it, treating the standard as a sort of recipe for how to get what they want. In contrast, an imprecise standard offers some scope for managers to convince the auditor or others of the validity of management's interpretation. Costly transaction structuring is not needed, and the paper suggests it might be less effective if it were done, perhaps because someone could argue a different interpretation of the imprecise standard that would challenge the structuring.

The theory about standards precision could have been developed further. One improvement would be in defining precision. The paper relates precision to rules, referring to precise and imprecise rules. Does this mean that precision implies form over substance, having rules to follow that obviate considering the underlying economic circumstance? This might imply that precision reduces earnings quality. A second improvement would be in exploring the connection between rules and the transaction structuring choice. If more rules means more precision, then would rule-based standards, or countries with them, be associated with relatively more transaction structuring than less "precise" principles-based standards? A third improvement would be in indicating how management's intent is implemented in earnings management attempts. Do managers who wish to manage earnings first search for imprecise standards that they can manipulate regardless of the underlying transactions, only thinking about transaction structuring as a last resort because it is costly? If standards are very precise, then does the earnings management goal lead to choosing transactions because they have a desired structure and so help achieve the goal, rather than because they are desirable transactions aside from the accounting (e.g. mergers undertaken for the purpose of creating income growth)?

The Auditor's Response

The paper views the auditor as opposed to earnings management (especially increases to current-year income), but responsive to the same standards precision and transaction structuring circumstances set out for client managers. The manager who wants to succeed in managing earnings need not calculate the auditor's likely response to an earnings management attempt. The manager should just choose the method that the manager would already prefer anyway (structuring transactions if standards are precise, not structuring them if the standards are imprecise) and the auditor will most likely not thwart the attempt. As the paper notes, this view of the auditor's response is behind the earlier theorizing about the manager's preferences: the managers prefer whichever method they think the auditor will accept. There is a flavor of auditor risk and power: the auditor has more trouble objecting if the standards are imprecise, or if a transaction has been structured to a precise standard. This flavor comes out also in the paper's expectation that the auditor's response will also depend on materiality and client size.

The auditor's role is passive, not prepared to counter a suitably designed management strategy. I would like to see the auditor's role more fully developed and more distinguished from management's intent. Such a separation might lead to more predictions. One might be that if the auditor's power is weaker when standards are imprecise, earnings management is more likely to involve imprecise standards than precise ones. (The paper did not hypothesize this, but the paper's Table 4 shows that 61 percent of the earnings management attempts (313/515) involved low-precision standards.) A second prediction, based on a negotiation or game theory view, might be that if the auditor is not in a strong position but wishes to influence the outcome, then the auditor would propose alternatives more acceptable to the auditor but still meeting management's intent to increase earnings. In such a case, the auditor could be seen as assisting in the earnings management, but that assistance would have been provided to thwart something less acceptable.

Hypotheses

Combining management's intention with the nature of applicable standards produces four earnings management scenarios to which the auditor will have to respond, shown in Table A.

TABLE A
Four Earnings Management Scenarios Used to Develop Hypotheses

	Standards Precision	
Managers' Intent	*Precise*	*Imprecise*
Increase current income	1	2
Decrease current income but allow increase in future	3	4

The paper's H1 predicts that standards precision leads to transaction structuring and imprecision to interpretation (no structuring). The hypothesis does not depend on whether the current-year income effect is an increase or a decrease. Therefore, the prediction is for transaction structuring across Table A's "precise" situations 1 and 3, and interpretation across "imprecise" situations 2 and 4. However, the paper's H4 predicts that interpretation is more likely when current income is being decreased. This prediction is for interpretation in situations 3 and 4. The predictions are therefore contradictory for situation 3.

In another prediction (H3), situations 1 and 2 will be more likely than situations 3 and 4. Hypothesis 5 additionally predicts that situation 4 is more likely than situation 3. (There is no frequency prediction for situation 1 vs. 2.) Combining H3 and H5, situation 3 should be the least frequent. The contradictory predictions of H1 and H4 noted above for situation 3 may therefore arise because the H1 effect is thought to come from a much larger frequency in situation 1 that overwhelms the tendency in situation 3 to go the other way.

The paper predicts (H2) that the auditor's response will follow managers' logic in H1—auditors will be less likely to oppose these earnings management attempts. The auditor, however, will not be so accommodating about management's preference for income-increasing earnings management and is more likely to oppose the frequent situations 1 and 2 than the infrequent situations 3 and 4 (H6). (The auditor also should seldom be troubled by the cut-and-dried and infrequent situation 3, but this is not hypothesized.) Across all situations, the auditor is also more likely to oppose earnings management attempts that are material (H7) or are proposed by smaller clients (H8).

Conclusion about Theory

The paper's theoretical support for its four planks is not extensive, flowing largely from definition and assertion. While the proposed relation between the precision of standards and transaction structuring seems plausible, the authors' conception of how this relation works could be developed better. In terms of capturing intent, the current-income effect leaves out other potential managerial motivations and is problematic if smoothing is part of the intent. However, though I would like more theoretical development, I do not think the paper's theoretical basis is wrong or that more development would substantially change a useful set of hypotheses.

II. THE STUDY'S OPERATIONALIZATION

The data come from a questionnaire completed by managers and partners, in which the respondents described earnings management attempts they had experienced and answered various questions about those attempts. This study had them do this three times, for three different experiences. (Results indicated that 515 earnings management attempts were collected from 253 audit partners and managers, an average of only two experiences.) Most

(80 percent) of the respondents were directed to a list of five potential earnings management areas (varied in order and content in the questionnaires sent out) and asked to choose their experiences from those, and 20 percent were not prompted. This retrospective, case-based method of gathering experiential data has naturalistic advantages and is useful in putting people's experiences in context. Leaving the respondents in control of the data leaves the researcher less able to discern clear relationships than in experimental studies, and usually leaves the researcher having to interpret important variables, as is the case here. The researcher's main impact is in the design of the questionnaire, and if the results are to be interpretable, then that design has to follow clear theory (Gibbins 2001). I concluded above that the theory is clear, if underdeveloped.

Intent

Managers' intent is one of the theory planks in this paper. The study approaches management intent partly by its definition of an earnings management attempt as influencing users and getting around standards to do so, but also examines intent in two ways. First, the questionnaire asked about managerial motivation, particularly regarding effects on current income. Second, the study separates the *attempt* to manage earnings from the financial statement *result*, proposing that the attempt reflects managers' intent better than the result, which is also affected by the auditor's response to the attempt.

The data are based on auditors' recall of earnings management attempts and their own responses. The evidence on managerial motivation thus depends on the accuracy of the auditors' understanding and recall. In the present study, this is a conceptual difficulty, in that the data gathered are all from the auditors; so, for example, auditor vulnerability in the face of imprecise standards is interpreted as managers' intent. We do not know if the managers in the cases reported by the auditors had such an intent, or realized that the auditor was vulnerable.

Earnings Management and Attempts to Manage Earnings

The paper takes it for granted that material attempts to manage earnings exist. Audit partners and managers responding to the questionnaire were asked to list experiences they had had where companies attempted to manage earnings. There was no opportunity to respond that earnings management attempts did not exist or had not been experienced or were immaterial, though the results show that some respondents appear to have disregarded the materiality instruction (more about this below).

Auditors who had not experienced material attempts presumably could indicate this by not responding to the questionnaire. Most auditors sent questionnaires did not respond. I do not doubt that auditors do experience attempts to manage earnings: for example, in our work on negotiation between auditors and clients (Gibbins et al. 2001), all the auditors reported having experienced such negotiations. However, the paper (in footnote 9) quotes one respondent as saying, "I am not now nor have been nor expect to be involved in any way with earnings management." The footnote is silent about whether any data from this auditor were included in the analysis, or whether there were other denials of experience with earnings management.

The questionnaire took an apparently broader approach to earnings management than the paper's view that earnings management is undesirable and the auditor must decide whether to allow it. The questionnaire defined such management as including "all efforts to influence accounting measurement, recognition and/or disclosure that are intended to [materially] affect the view of an entity obtained by a person using the financial statements," and earnings management attempts to include "all experiences where the company attempted to manage earnings, whether or not the final audited financial statements appeared as the company initially desired" (see the Appendix to the paper).

These definitions and the rest of the questionnaire put no explicit negative connotation on earnings management, so we have to assume the auditors did, in fact, view it negatively. As some discussion at the Conference showed, we also must take on some faith the connection between reported attempts to management earnings and reduced quality of earnings, because by the questionnaire's definition, any attempt to influence the financial statements is earnings management, regardless of motivation. Managing earnings to make them a better measure of performance, or to make them more prudent or conservative would fit the questionnaire's definition, if not the paper's view of earnings management. The second paragraph of the questionnaire (see the paper's Appendix) specified the two anticipated earnings management mechanisms (structuring transactions and interpreting standards), and may have created the authors' desired negative view by using the phrases "gets around a standard" and "aggressively interpreting the wording of a standard."

Variables

Four variables were central to the analysis. The first three were developed by coding the auditor respondents' narratives. Precision was coded 1 if precise and 0 if imprecise, with some bias toward coding multi-element standards as imprecise if at least one of the elements was imprecise. Structuring was coded 1 if present and 0 if not (no coding was done of the complement, interpreting standards). Current year's income effect was coded 1 if current income was increased, −1 if it was decreased, and 0 if there was no clear effect. All these codings converted potentially complex and subtle variables into simple categories. Thus, the analysis is simplified to match the simplified theory.

There was some discussion at the Conference about how to distinguish between transactions chosen (or structured) for business reasons and those chosen (or structured) simply to achieve the desired accounting result. The coding did not attempt to distinguish between these, so some of the structured attempts may have been for good business reasons. However, the paper's Table 4 indicates that there were relatively few transactions-structuring earnings management attempts (67 of 515), so any ambiguity about whether the structuring was for good business reasons would have little effect on the results.

This conversion of the data into categories was done by coders who had no knowledge of the hypotheses. Inter-coder reliability was checked and the analyses were run with individual coders' classifications as well as the consensus used in the paper's tables. Substantial confidence in the role of the three coded variables in the analysis seems warranted.

The fourth variable, how the earnings management attempt was resolved, was collected directly from the respondents, who were asked to choose among these four categories of outcomes as treatments in the financial statements (see the paper's footnote 16; numbers below are from that footnote):

- "treated as the company originally desired, because they demonstrated they were complying with GAAP" (110 cases); or
- "treated as the company originally desired, because there was no convincing evidence that the company's position was incorrect" (88 cases); or
- "treated as the company originally desired, for some other reason (please explain)" (91 cases); or
- "treated other than how the company originally desired (please explain)" (226 cases).

In accordance with the simplified theory that the auditor can allow or disallow an earnings management attempt, these four categories were collapsed into two for the results analysis. The authors interpret the last category as adjusted (required by the auditor), and the first three as not adjusted. Other solutions, such as negotiation (Gibbins et al. 2001; Beattie et al. 2001) are not considered. If the auditor insisted on "winning" the dispute (which the negotiation literature calls a "distributive" solution), then the auditor would choose the fourth outcome. But if the auditor and client negotiated a solution not originally anticipated (e.g., an "integrative" solution in which both parties gain), then the auditor would also have to choose the fourth outcome. My view is that the binary variable is not clearly interpretable because of the variety of outcomes captured in the fourth category.

III. THE RESULTS

The paper's theory could have made more of incentive issues, auditor responses such as negotiation, and the costliness of transaction structuring relative to exerting pressure on the auditor. But the questionnaire operationalized the simplified theory, so its results should be interpreted against its theory.

Response Context

The purpose of the study's questionnaire is to gather experiential data, so the nature of the respondent group and their range of experiences is central to the results. The study's sample was large and senior (audit partners and managers), but the 16 percent response rate was poor. Though the paper goes to some length to defend the results as valid anyway, the threat of selective responses, such as by those who had time on their hands or did not have anything embarrassing to report, is real. Further, as the respondents provided an average of two earnings management attempts each, though three were requested, there must have been considerable fatigue, or else the auditors did not have as much experience with earnings management attempts as had been anticipated.

The authors mailed 532 packages, requesting that three auditors (a partner and two managers) give three experiences each, for a potential of 4,788 experiences. They received 515 usable experiences, or only 11 percent of the potential. This reinforces the point that the results cannot be considered descriptive, but only relevant to the theory testable from the responses received. The data cannot tell us whether earnings management attempts are a serious problem, how often auditors experience them, or how they feel about them, because there are so many possible reasons for the 89 percent of missing data.

As was discussed frequently at the Conference and is addressed by the authors, the data from such a questionnaire are retrospective and must be treated with some caution. This caveat applies to most studies regardless of method, so I do not single this study out in that regard. For this study, the caution relates to response biases of various kinds, memory defects, fatigue, and failure to follow instructions. Some of the worry about retrospective data comes from the effects of such potential problems on the descriptive validity of the results. This problem is unfortunately exacerbated by the study's use of only one public accounting firm as the source of data. While I have no reason to doubt that the simplified theory set out in the paper would apply to any firm, my conclusion is that the paper's descriptive value is minimal. The paper's Tables 2 and 3 should be taken as suggestive, not descriptive. Paradoxically perhaps, lack of descriptive value mitigates the concern about retrospective data: biases that reduce descriptive value may cancel out in theory-relevant comparisons within the data. For example, if earnings management attempts are generally over- or underreported by the respondents, then comparisons of the number of attempts that increase or decrease earnings will be valid relatively, as long as there is little likelihood that the data biases are correlated with either increase or decrease attempts.

Materiality

Respondents were asked to describe material earnings management attempts, but many of the 515 reported instances were described by the auditors themselves as immaterial. Combining the 66 of the "unadjusted" earnings management attempts reported in footnote 16 with the 22 of the "adjusted" attempts reported in the body of the paper under the heading "Materiality and Client Size" gives 88 immaterial attempts, 17 percent of the total experiences. According to Table 4, there were only 67 of the hypothesized transaction-structured earnings management attempts, so the 88 immaterial attempts are numerous in relation to the theory about such attempts. Footnote 7 says that 377 of the 515 attempts were identified by the auditors as material, and that using only those in the analysis leaves the results "very similar," which is reassuring. The respondents may, as the paper argues, have reconsidered materiality as they described the experience, but if so, the changeable materiality criterion may reflect a retrospection bias in that respondents were not initially able to follow the instruction to report on material experiences.

Hypotheses

Subject to my previous comments and the authors' careful delineation of threats to validity, the results support the hypotheses well. Reworking the paper's Tables 4 and 6 into the same tabular format I used earlier to discuss the hypotheses produces Table B (TS = transaction structuring; NTS = nonstructuring).

Most of the earnings management attempts were income-increasing, as hypothesized. However, they were only 53 percent of the total sample (272/515), not a clear majority as might be expected from H3's prediction that managers are more likely to attempt to increase current income. The test of H3 used only the 431 increases and decreases in Table B (the paper's Table 6), and tested their difference in the presence of the precision and structuring variables. This is not a clear test of H3's simple assertion that income-increasing attempts are more likely, because it separates all the interactions between cell frequencies and isolates only the residual income effect. The test is therefore whether income increasing is more frequent than income decreasing, after accounting for precision and structuring effects. I like the test, but therefore would have liked H3 to have been stated and supported as being the more restricted relation tested.

TABLE B
Results from the Paper's Tables 4 and 6 Organized According to Table A

	Standards Precision						
	Precise			*Imprecise*			
Managers' Intent	*NTS*	*TS*	*Total*	*NTS*	*TS*	*Total*	*Sum*
Increase current income	91	29	120	136	16	152	272
Decrease current income	42	4	46	113	0	113	159
Sums (Table 6)	133	33	166	249	16	265	431
Income effect unclear	23	13	36	43	5	48	84
Sums (Table 4)	156	46	202	292	21	313	515

The frequency of transaction structuring was low, virtually absent in the case of decreases in current income. These results amount to more reasons for better theorizing about cases other than income-increasing earnings management attempts. The hypothesized interactions of earnings management mechanism and standards precision appeared, albeit limited by the small frequencies of transaction structuring, and by the similarity of the pattern of nonstructured vs. structured earnings management attempts in all four hypothesized situations, with nonstructured dominating in every one. As I suggested earlier, this dominance of nonstructured attempts might have been predicted from the paper's theory.

IV. CONCLUSION

The paper is an interesting exploration of an important question: if managers intend to manipulate earnings, then how does that intent work through to the financial statements, past the barriers of implementation and auditor objections? The results fit the authors' expectations as based on their theory and are intuitively appealing, though not complimentary to auditors in their suggestion that if managers know where the auditor's barrier is weak, then they can arrange their earnings management mechanisms to succeed. I have argued that the paper's results should be interpreted given some weaknesses in theory, problems in operationalizing the theory, lack of descriptive validity, and some data problems. The paper indicates that the authors are generally aware of these problems, and I look forward to further research that will address them and further illuminate management's intent.

REFERENCES

Beattie, V., S. Fearnley, and R. Brandt. 2001. *Behind Closed Doors*. Sponsored by the Institute of Chartered Accountants in England and Wales. Basingstoke, Hamshire, U.K.: Palgrave.

Gibbins, M., J. Newton. 1994. An empirical examination of complex accountability in public accounting. *Journal of Accounting Research* 32 (2): 165–186.

———. 2001. Incorporating context into the study of judgment and expertise in public accounting. *International Journal of Auditing* 5 (3): 225–236.

———, S. Salterio, and A. Webb. 2001. Evidence about auditor-client management negotiation concerning client's financial reporting. *Journal of Accounting Research* 39 (3): 535–563.

———, and K. Trotman. 2002. Audit review: Managers' interpersonal expectations and conduct of the review. *Contemporary Accounting Research* 19 (3).

Nelson, M. W., J. A. Elliott, and R. L. Tarpley. 2002. Evidence from auditors about managers' and auditors earnings management decisions. *The Accounting Review* (Supplement): 175–202.

Schipper, K. 1989. Commentary on earnings management. *Accounting Horizons* 3 (December): 91–102.

EDITORIAL POLICY AND STYLE INFORMATION

EDITORIAL POLICY

According to the policies set by the Publications Committee (which were endorsed by the Executive Committee and were published in the *Accounting Education News*, June 1987), *The Accounting Review* "should be viewed as the premier journal for publishing articles reporting the results of accounting research and explaining and illustrating related research methodology. The scope of acceptable articles should embrace any research methodology and any accounting-related subject, as long as the articles meet the standards established for publication in the journal....No special sections should be necessary. The primary, but not exclusive, audience should be—as it is now—academicians, graduate students, and others interested in accounting research."

The primary criterion for publication in *The Accounting Review* is the significance of the contribution an article makes to the literature.

The efficiency and effectiveness of the editorial review process is critically dependent upon the actions of both authors submitting papers and the reviewers. Authors accept the responsibility of preparing research papers at a level suitable for evaluation by independent reviewers. Such preparation, therefore, should include subjecting the manuscript to critique by colleagues and others and revising it accordingly prior to submission. The review process is not to be used as a means of obtaining feedback at early stages of developing the research.

Reviewers and associate editors are responsible for providing critically constructive and prompt evaluations of submitted research papers based on the significance of their contribution and on the rigor of analysis and presentation. Associate editors also make editorial recommendations to the editor.

MANUSCRIPT PREPARATION AND STYLE

The Accounting Review's manuscript preparation guidelines follow (with a slight modification) the B-format of *The Chicago Manual of Style* (14th ed.; University of Chicago Press). Another helpful guide to usage and style is *The Elements of Style*, by William Strunk, Jr., and E. B. White (Macmillan). Spelling follows *Webster's Collegiate Dictionary.*

FORMAT

1. All manuscripts should be typed in 12-point font on one side of 8½ × 11" good quality paper and be double spaced, except for indented quotations.
2. Manuscripts should be as concise as the subject and research method permit, generally not to exceed 7,000 words.
3. Margins of at least one inch from top, bottom, and sides should facilitate editing and duplication.
4. To promote anonymous review, authors should not identify themselves directly or indirectly in their papers or in experimental test instruments included with the submission. Single authors should not use the editorial "we."
5. A cover page should show the title of the paper, the author's name, title and affiliation, email address, any acknowledgments, and a footnote indicating whether the author would be willing to share the data (see last paragraph in this statement).

Pagination: All pages, including tables, appendices and references, should be serially numbered. Major sections should be numbered in Roman numerals. Subsections should not be numbered.

Numbers: Spell out numbers from one to ten, except when used in tables and lists, and when used with mathematical, statistical, scientific, or technical units and quantities, such as distances, weights and measures. For example: *three days; 3 kilometers; 30 years.* All other numbers are expressed numerically.

Percentages and Decimal Fractions: In nontechnical copy use the word percent in the text.

Hyphens: Use a hyphen to join unit modifiers or to clarify usage. For example: *a well-presented analysis; re-form.* See *Webster's* for correct usage.

Keywords: The abstract is to be followed by four keywords that will assist in indexing the paper.

ABSTRACT/INTRODUCTION

An Abstract of about 100 words should be presented on a separate page immediately preceding the text. The Abstract should concisely inform the reader of the manuscript's topic, its methods, and its findings.

Keywords and the Data Availability statements should follow the Abstract. The text of the paper should start with a section labeled "I. Introduction," which provides more details about the paper's purpose, motivation, methodology, and findings. Both the Abstract and the introduction should be relatively nontechnical, yet clear enough for an informed reader to understand the manuscript's contribution. The manuscript's title, but neither the author's name nor other identification designations, should appear on the Abstract page.

TABLES AND FIGURES

The author should note the following general requirements:

1. Each table and figure (graphic) should appear on a separate page and should be placed at the end of the text. Each should bear an Arabic number and a complete title indicating the exact contents of the table or figure. Tables and figures should define each variable. The titles and definitions should be sufficiently detailed to enable the reader to interpret the tables and figures without reference to the text.
2. A reference to each graphic should be made in the text.
3. The author should indicate by marginal notation where each graphic should be inserted in the text.
4. Graphics should be reasonably interpreted without reference to the text.
5. Source lines and notes should be included as necessary.

Equations: Equations should be numbered in parentheses flush with the right-hand margin.

DOCUMENTATION

Citations: Work cited should use the "author-date system" keyed to a list of works in the reference list (see below). Authors should make an effort to include the relevant page numbers in the cited works.

1. In the text, works are cited as follows: authors' last name and date, without comma, in parentheses: for example, (Jones 1987); with two authors: (Jones and Freeman 1973); with more than two: (Jones et al. 1985); with more than one source cited together (Jones 1987; Freeman 1986); with two or more works by one author: (Jones 1985, 1987).
2. Unless confusion would result, do not use "p." or "pp." before page numbers: for example, (Jones 1987, 115).
3. When the reference list contains more than one work of an author published in the same year, the suffix a, b, etc. follows the date in the text citation: for example, (Jones 1987a) or (Jones 1987a; Freeman 1985b).
4. If an author's name is mentioned in the text, it need not be repeated in the citation; for example, "Jones (1987, 115) says...."
5. Citations to institutional works should use acronyms or short titles where practicable; for example, (AAA ASOBAT 1966); (AICPA Cohen Commission Report 1977). Where brief, the full title of an institutional work might be shown in a citation: for example, (ICAEW *The Corporate Report* 1975).
6. If the manuscript refers to statutes, legal treatises or court cases, citations acceptable in law reviews should be used.

Reference List: Every manuscript must include a list of references containing only those works cited. Each entry should contain all data necessary for unambiguous identification. With the author-date system, use the following format recommended by *The Chicago Manual*:

1. Arrange citations in alphabetical order according to surname of the first author or the name of the institution responsible for the citation.
2. Use author's initials instead of proper names.
3. Date of publication should be placed immediately after author's name.
4. Titles of journals should not be abbreviated.
5. Multiple works by the same author(s) in the same year are distinguished by letters after the date.
6. Inclusive page numbers are treated as recommended in *Chicago Manual* section 8.67.

Sample entries are as follows:

American Accounting Association, Committee on Concepts and Standards for External Financial Reports. 1977. *Statement on Accounting Theory and Theory Acceptance*. Sarasota, FL: AAA.

Demski, J. S., and D. E. M. Sappington. 1989. Hierarchical structure and responsibility accounting. *Journal of Accounting Research* 27 (Spring): 40–58.

Dye, R., B. Balachandran, and R. Magee. 1989. Contingent fees for audit firms. Working paper, Northwestern University, Evanston, IL.

Fabozzi, F., and I. Pollack, eds. 1987. *The Handbook of Fixed Income Securities*. 2nd edition. Homewood, IL: Dow Jones-Irwin.
Kahneman, D., P. Slovic, and A. Tversky, eds. 1982. *Judgment Under Uncertainty: Heuristics and Biases*. Cambridge, U.K.: Cambridge University Press.
Porcano, T. M. 1984a. Distributive justice and tax policy. *The Accounting Review* 59 (October): 619–636.
———. 1984b. The perceived effects of tax policy on corporate investment intentions. *The Journal of the American Taxation Association* 6 (Fall): 7–19.
Shaw, W. H. 1985. Empirical evidence on the market impact of the safe harbor leasing law. Ph.D. dissertation, The University of Texas at Austin.
Sherman, T. M., ed. 1984. *Conceptual Framework for Financial Accounting*. Cambridge, MA: Harvard Business School.

Footnotes: Footnotes are not used for documentation. Textual footnotes should be used only for extensions and useful excursions of information that if included in the body of the text might disrupt its continuity. Footnotes should be consecutively numbered throughout the manuscript with superscript Arabic numerals. Footnote text should be doubled-spaced and placed at the end of the article.

SUBMISSION OF MANUSCRIPTS

Authors should note the following guidelines for submitting manuscripts:

1. Manuscripts currently under consideration by another journal or publisher should not be submitted. The author must state that the work is not submitted or published elsewhere.
2. In the case of manuscripts reporting on field surveys or experiments, four copies of the instrument (questionnaire, case, interview plan or the like) should be submitted. Information that might identify the author(s) must be deleted from the instrument.
3. The manuscript should be sent in an email as a Word or PDF attachment, with the title page in a separate file. ONE paper copy should be submitted with a check in U.S. funds for $75.00 for members or $100.00 for nonmembers of the AAA made payable to the American Accounting Association. The submission fee is nonrefundable. If preferable, the manuscript can be sent on a disk instead of as an email attachment, along with the one paper copy and the check.
4. The author should retain a copy of the paper.
5. Revisions must be submitted within 12 months from request, otherwise they will be considered new submissions.

COMMENTS

Comments on articles previously published in *The Accounting Review* will be reviewed (anonymously) by two reviewers in sequence. The first reviewer will be the author of the original article being subjected to critique. If substance permits, a suitably revised comment will be sent to a second reviewer to determine its publishability in *The Accounting Review*. If a comment is accepted for publication, the original author will be invited to reply. All other editorial requirements, as enumerated above, also apply to proposed comments.

POLICY ON REPRODUCTION

An objective of *The Accounting Review* is to promote the wide dissemination of the results of systematic scholarly inquiries into the broad field of accounting.

Permission is hereby granted to reproduce any of the contents of the *Review* for use in courses of instruction, as long as the source and American Accounting Association copyright are indicated in any such reproductions.

Written application must be made to the American Accounting Association, 5717 Bessie Drive, Sarasota, FL 34233-2399, for permission to reproduce any of the contents of the *Review* for use in other than courses of instruction—e.g., inclusion in books of readings or in any other publications intended for general distribution. In consideration for the grant of permission by the *Review* in such instances, the applicant must notify the author(s) in writing of the intended use to be made of each reproduction. Normally, the *Review* will not assess a charge for the waiver of copyright.

Except where otherwise noted in articles, the copyright interest has been transferred to the American Accounting Association. Where the author(s) has (have) not transferred the copyright to the Association, applicants must seek permission to reproduce (for all purposes) directly from the author(s).

POLICY ON DATA AVAILABILITY

The following policy has been adopted by the Executive Committee in its April 1989 meeting.

"An objective of (*The Accounting Review, Accounting Horizons, Issues in Accounting Education*) is to provide the widest possible dissemination of knowledge based on systematic scholarly inquiries into accounting as a field of professional research, and educational activity. As part of this process, authors are encouraged to make their data available for use by others in extending or replicating results reported in their articles. Authors of articles which report data dependent results should footnote the status of data availability and, when pertinent, this should be accompanied by information on how the data may be obtained."